

University of Groningen

Comparative and functional genomics of lactococci

Kok, Jan; Buist, Girbe; Zomer, Aldert L.; Hijum, Sacha A.F.T. van; Kuipers, Oscar P.

Published in:
FEMS Microbiology Reviews

DOI:
[10.1016/j.fmrre.2005.04.004](https://doi.org/10.1016/j.fmrre.2005.04.004)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2005

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Kok, J., Buist, G., Zomer, A. L., Hijum, S. A. F. T. V., & Kuipers, O. P. (2005). Comparative and functional genomics of lactococci. *FEMS Microbiology Reviews*, 29(3), 411-433.
<https://doi.org/10.1016/j.fmrre.2005.04.004>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Comparative and functional genomics of lactococci

Jan Kok *, Girbe Buist, Aldert L. Zomer, Sacha A.F.T. van Hijum, Oscar P. Kuipers

Department of Molecular Genetics, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen,
Kerklaan 30, 9751 NN Haren, The Netherlands

Received 10 January 2005; accepted 22 April 2005

First published online 28 August 2005

Abstract

Whole-genome nucleotide sequencing has revolutionized the genetic, biochemical and molecular biology research on bacteria and indeed, many higher organisms. The genome sequences of the strains of two subspecies of *Lactococcus lactis*, *L. lactis* subsp. *lactis* and *L. lactis* subsp. *cremoris*, have been determined. These genomic sequences have permitted two important new approaches to be applied in the research of *L. lactis*. The analysis of the regulation of expression of all genes under specific circumstances at a given point in time is now possible by DNA microarray technology. The elucidation of the full protein complement of the organism as a function of intrinsic or external factors has been made possible by high-throughput protein identification and analysis techniques combined with the gene-derived know-how of the total protein encoding capacity of the genome. These techniques from the genomics arena, transcriptomics and proteomics, have been recently implemented in the study of various aspects of growth and functioning of *L. lactis*. In this paper we discuss a number of similarities and differences between the two lactococcal genome sequences and review the current status of genomics research in *L. lactis*. We also propose future directions with respect to both answering fundamental questions more quickly and more completely, as well as opening new avenues for biotechnological applications.

© 2005 Federation of European Microbiological Societies. Published by Elsevier B.V. All rights reserved.

Keywords: *Lactococcus lactis*; Genome; DNA microarray; Transcriptomics; Proteomics

Contents

1. Introduction	412
2. Genome plasticity in <i>L. lactis</i>	412
3. The genetic content of <i>L. lactis</i>	413
3.1. IS elements and prophages	413
3.2. Plasmid integrations in <i>L. lactis</i> MG1363	414
3.3. Some of the genetic differences between <i>L. lactis</i> MG1363 and <i>L. lactis</i> IL1403	415
3.4. IS element-derived mutations in <i>L. lactis</i>	416
4. Global analyses of <i>L. lactis</i> – transcriptomics	416
4.1. Transcriptome analyses in <i>L. lactis</i> using partial DNA arrays	417
4.2. Nitrogen metabolism	418
4.3. Identification of a novel glucose-PTS system in <i>L. lactis</i> by transcriptome analysis	422
4.4. Fructose metabolism in <i>L. lactis</i>	422

* Corresponding author. Fax: +31 50 363 2348.

E-mail address: jan.kok@rug.nl (J. Kok).

4.5.	Transcriptome analysis reveals nisin-resistance mechanisms of <i>L. lactis</i>	423
4.6.	<i>L. lactis</i> and food safety	424
5.	Genomic diversity analysis using comparative genome hybridization	425
5.1.	Comparing gene expression profiles across organisms	425
6.	Global analyses of <i>L. lactis</i> – proteomics	426
6.1.	Generation of <i>L. lactis</i> proteome reference maps	427
6.2.	Quantitative proteomics	427
7.	Concluding remarks	428
	Acknowledgments	430
	References	430

1. Introduction

Lactococci have a long history of use in milk fermentations, from small-scale traditional operations on the farm or in the family home to ever-increasing industrial scale processes. The consequent need for more robust, efficient and fine-tuned manufacturing practice has led to a sharp increase in fundamental and applied research on the bacterial species involved, *Lactococcus lactis*. Tremendous advances have been made in unravelling the genetics and molecular biology of these economically very important microorganisms. Many tools and techniques have been developed to genetically dissect *L. lactis*, such as plasmid cloning vectors and transformation methods, various (inducible) gene expression vectors and methodologies to introduce any mutation in specified regions of the genome of *L. lactis*. Thus, this organism has become the paradigm of lactic acid bacteria (LAB) and the second-best studied gram-positive bacterium, only surpassed by *Bacillus subtilis*. The increased knowledge of the genetics, physiology and molecular biology of *L. lactis* has allowed deep understanding of the traits that are of utmost importance for the industrial processes in which the organism is used. The route, e.g., by which it takes up and metabolizes lactose, leading to the rapid pH-drop during milk fermentation, its ability to degrade proteins and peptides necessary for proper cheese flavour development or its capacity to produce diacetyl have all been extensively studied. It has also allowed us to learn more about the organisms' bacteriophages and to prevent the damaging effects these viruses can have on the industrial processes. Entirely new directions of studying and modifying *L. lactis* are possible with the advanced genetic engineering techniques that are currently at the disposal of academic and industrial scientists. Some recent highlights that may exemplify the breath of possibilities are (1) the wilful rerouting of metabolism in *L. lactis* in such a way that it produces alanine [1]; (2) the modification of *L. lactis* to overproduce the vitamin folate to such

an extent that the strain could, in principle, be used in functional-food formulations [2]; (3) the heterologous secretion of human interleukin-10 by *L. lactis* and the effectiveness of live lactococcal cells in treating chemical-induced colitis in mice [3]. The latter may be an illustration of the potential to use this bacterium for oral delivery of therapeutic molecules [4,5]. On another level, the relatively simple metabolic activity, low proteolytic capacity and genetic amenability make it an ideal organism to (over)express foreign (membrane) proteins [6,7] or to clone and study pathogenicity factors of related (low-GC gram-positive) pathogens [8].

The scope of research possibilities, and thus of studying existing or novel application potential of *L. lactis*, has recently expanded through the elucidation of the nucleotide sequence of the genome of *L. lactis*. This has allowed devising full-genome DNA-microarrays for the two laboratory strains *L. lactis* subsp. *lactis* IL1403 and *L. lactis* subsp. *cremoris* MG1363. Together with developments in high-throughput protein analysis these genomics techniques will deepen our understanding of the functionality of *L. lactis* by offering a global view of how the organism reacts to changes in its internal and external environment. They will prove to be of eminent importance for the development and/or selection of improved strains for the existing and novel (food) applications.

The aim of this paper is to summarize the current knowledge on functional genomics of lactococci and to examine the future perspectives of genome mining and experimental genomics research on *L. lactis*.

2. Genome plasticity in *L. lactis*

Bacteria can be divided into two large groups with regard to the stability of their genomes [9] and references therein). One group consists of bacterial species of which the genomic organization is highly variable, such as *Bacillus cereus* and *Pseudomonas aeruginosa*. *L. lactis*,

on the other hand, belongs to the group of bacteria with relatively stable genomes, which includes the paradigm species *B. subtilis* and *Escherichia coli*. Nevertheless, genetic rearrangements have been observed early on when the two laboratory strains *L. lactis* ssp. *cremoris* MG1363 and *L. lactis* ssp. *lactis* IL1403 and several other strains of *L. lactis* were compared [10–12]. *L. lactis* MG1363 is a plasmid-free derivative of *L. lactis* ssp. *cremoris* NCDO712, the progenitor of the NCDO712 family of strains, which differ with respect to plasmid or prophage DNA, the location of the conjugative sex factor (either absent or present as a low-copy-number plasmid or inserted into the chromosome), or the presence of a region encompassing the genes specifying the oligopeptide transport system Opp [13–18]. In-depth analyses using restriction enzyme digestions in combination with pulsed-field gel electrophoresis revealed the presence of five macro-restriction fragment length polymorphisms (macro-RFLPs) [19]. Using a physical map of the chromosome of *L. lactis* MG1363, the five macro-RFLPs could be assigned to certain genetic events at five particular regions in the MG1363 genome [19] (see Fig. 2 from Campo et al. [9] for a comprehensive picture of the rearrangements). Homologous recombination between two copies of IS905 was shown to be the cause of a large inversion of half of the chromosome in *L. lactis* NCDO763 [20]. Although several of the strains of the NCDO712 family were reported to be free of prophages, phage-related elements were present in their chromosomes. In fact all strains carried a defective phage-related region [20], while seven out of nine strains investigated carried an additional 40-kb excisable uncharacterized prophage [19]. The *opp-pepO* region is part of an approximately 25-kb transposon-like structure that is located on a plasmid in certain strains, in the chromosome of others, or absent altogether. This structure is present in a region of the genome of *L. lactis* MG1363 that is prone to DNA insertion and recombination (see below). Thus, as concluded by Campo et al. [9], two isogenic *L. lactis* laboratory strains can differ with respect to genomic content by more than 5% by the presence of large excisable elements such as prophages and transposon-like structures.

3. The genetic content of *L. lactis*

The genome sequences of *L. lactis* ssp. *lactis* IL1403 [21] and *L. lactis* ssp. *cremoris* MG1363 (Zomer, A., Wegmann, U., O'Connell-Motherway, M., Buist, G., Goesmann, A., de Jong, A., Gasson, M., Shearman, C., Kuipers, O., van Sinderen, D. and Kok, J., unpublished) have been sequenced to completion, while that of the *L. lactis* ssp. *cremoris* strain SK11 is nearing completion [22] (www.jgi.doe.gov). A comparison of the genetic content of the two fully sequenced strains is given

in Table 1. The chromosome of *L. lactis* MG1363 (2.53 Mb) is 160 kb larger than that of *L. lactis* IL1403 (2.37 Mb) and has an average GC content of 35.8%. Consequently, because of its larger size, the MG1363 genome encodes more proteins: while the IL1403 genome specifies 2270 predicted proteins larger than 60 amino acid residues, that of MG1363 carries at least 2459 protein-encoding genes larger than 60 amino acid residues, with an average size of 872 bp. This results in a coding density of approximately 85% for both complete sequenced *Lactococcus* genomes, which is similar to what has been observed for genomes of related organisms. Six hundred and eighty-four of the deduced amino acid sequences of *L. lactis* MG1363 are not present in *L. lactis* IL1403; of these proteins 199 have been described before in other lactococcal strains, 281 are similar to non-lactococcal proteins while the remaining 204 proteins did not show any homology to other proteins in data bases (based on blast searches against the non-redundant protein database with an *E*-value cut-off of 10^{-20}). Similarly, 367 of the deduced amino acid sequences of *L. lactis* IL1403 are not present in *L. lactis* MG1363 (based on blast searches against the *L. lactis* MG1363 proteins with an *E*-value cut-off of 10^{-20}). Using automated COG function prediction a putative function could be assigned to 1887 of the *L. lactis* MG1363 open reading frames (ORFs) (77%), which is similar to the results obtained for IL1403 (1789 of the 2270 ORFs (79%)). A comparison of the locations of the ORFs in the genome of *L. lactis* MG1363 with those of *L. lactis* IL1403 illustrates the presence of the large chromosomal inversion described above (Fig. 1).

3.1. IS elements and prophages

The number of IS elements in *L. lactis* MG1363 was determined by aligning all putative proteins specified by the MG1363 genome on the non-redundant protein database provided by NCBI using the BLAST algorithm [23]. *L. lactis* MG1363 contains at least 10 different IS elements whereas *L. lactis* IL1403 carries six different IS elements (Table 2) [21]. The IS elements in *L. lactis* MG1363 comprise a total of 66 kb of DNA. The IS elements of *L. lactis* IL1403 are unequally distributed: the

Table 1
Comparison of characteristics of nucleotide sequences of *L. lactis* IL1403 and *L. lactis* MG1363

Strain	<i>L. lactis</i> MG1363	<i>L. lactis</i> IL1403
Size (Mb)	2.53	2.37
Nr. of ORFs	~2500	~2310
Total GC %	35.8	35.4
GC % ORFs	36.7	36.1
Nr. phage genes	~200	293
Phage DNA (kb)	134	293
IS elements	92	52

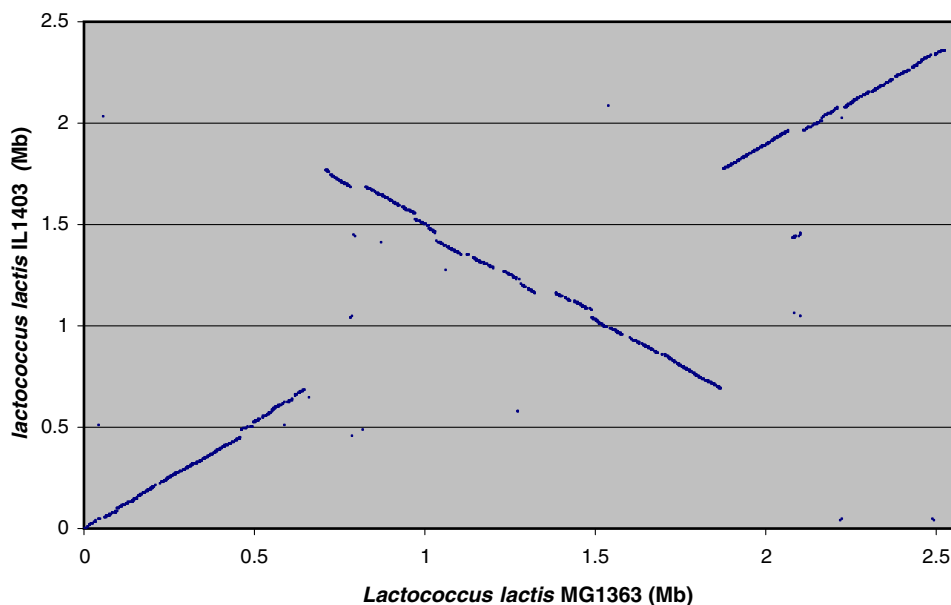


Fig. 1. Dot plot of the positions of homologous ORFs of *L. lactis* MG1363 and *L. lactis* IL1403. Homology was determined using BLASTN at a cut-off of $1e^{-20}$. Repetitive sequences in both strains were removed prior to the BLASTN searches by the repeat masking module in *Projector 2* [121].

Table 2
Comparison of IS elements of *L. lactis* MG1363 and *L. lactis* IL1403

IS element (-like)	MG1363	IL1403
IS904	9	9
IS1077	9	7
IS905	16	1
IS981	16	10
IS982	2	1
IS983	0	15
IS1076	10	9
<i>tmpA</i> homolog (<i>S. pyogenes</i>) NP_664202	7	0
Listeria transposase NP_569202	8	0
IS946/ISS1	3	0
Unknown	12	0
Total number of ISs	92	52

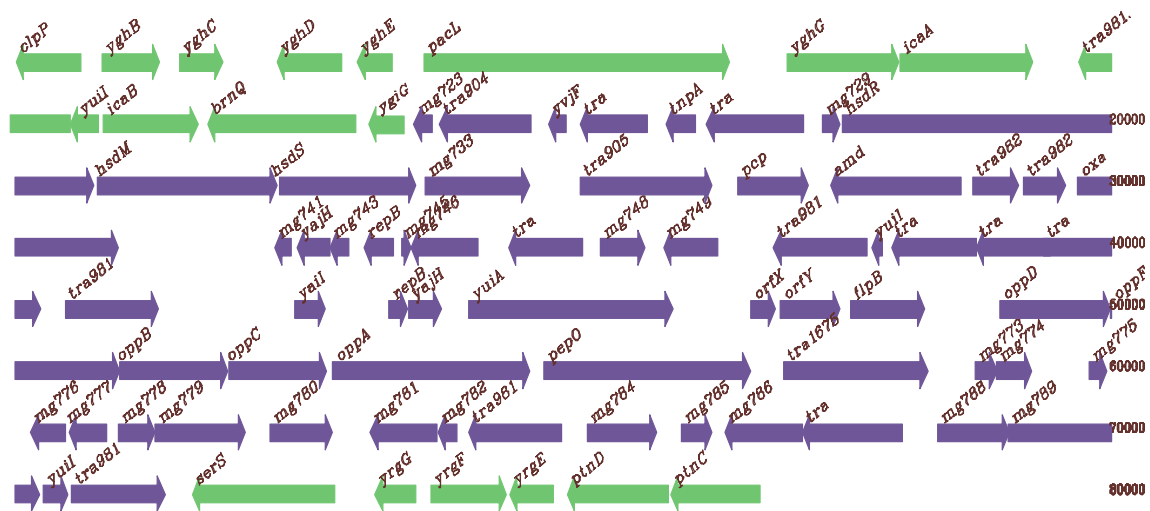
seven copies of IS1077 occupy a region of 1.0 Mb while IS983 is present on 1.5 Mb DNA fragment. Both regions overlap by approximately 150 kb. Together with the fact that the 10 copies of another element, IS981, are distributed over the entire genome, these observations suggest that *L. lactis* IL1403 is the product of a fusion of two large chromosomal DNA fragments [21].

Quite a number of the ISs in *L. lactis* MG1363 are flanked by sequences not present in *L. lactis* IL1403, such as, e.g., the conjugative sex factor. The presence of this piece of DNA, which can be seen sometimes as a low-copy-number extra-chromosomal element, enables *L. lactis* MG1363 to transfer DNA at a high frequency, including chromosomal sequences [24,25]. Although *L. lactis* MG1363 has gone through a prophage curing scheme, three complete bacteriophage genomes and three regions containing phage remnants are

present in its chromosome. One of the complete phages is homologous to the lactococcal phage bIL310, which is present in *L. lactis* IL1403, and another to lactococcal phage T712. All complete phages contain putative transposase genes. These prophage regions cover a total of at least 134 kb of the genome of *L. lactis* MG1363. Thus, although this amount of prophage DNA is less than in *L. lactis* IL1403, which has 175 kb of prophage DNA [21], the genome of *L. lactis* MG1363 is certainly not free of prophage DNA.

3.2. Plasmid integrations in *L. lactis* MG1363

Major differences in the gene content of *L. lactis* IL1403 and *L. lactis* MG1363 are caused by the integration of mobile elements, one of which is the sex factor mentioned above. Another known integration in the genome of *L. lactis* MG1363 is the *opp-pepO* region, specifying *a.o.* the oligopeptidase PepO and the oligopeptide uptake system Opp [17,26], providing the strain with a second *opp* system. Analysis of the surroundings of *opp-pepO* suggests that a much larger integration or multiple integrations took place in this area. This region, which is not present in IL1403, has a total size of 56 kb and contains multiple IS elements, genes for invertases/resolvases and remnants of genes of plasmid replication proteins (Fig. 2). Furthermore, it contains the *hsdR/M/S* genes encoding a restriction modification system often found on plasmids, and the gene for pyrrolidone carboxyl peptidase, *pcp*, which has also been suggested to be of plasmid origin (Buist, G., Haandrikman, A.J., Benus, G., Feenstra, B., Venema, G. and Kok, J., unpublished). Various other hypothetical ORFs, which were



L. lactis MG1363 does not contain *cit*, the operon encoding the enzymes for citrate utilization. Examples of other genes that are absent in comparison to *L. lactis* IL1403 are *otcA* and *arcC3*. The *arcC3* gene encodes a carbamate kinase. Two paralogs of *arcC3*, *arcC1* and *arcC2*, are present in *L. lactis* MG1363. OtcA (ornithine carbamoyltransferase (EC 2.1.3.3)) is involved in the conversion of ornithine to citrulline. However, a recent reinterpretation of genome annotations regarding carbamoyltransferases suggests that lactococcal OtcA might in fact be a putrescine carbamoyltransferase [31]. *L. lactis* IL1403 *otcA-yrfD-yrfC-arcC3* encode a putative catabolic agmatine deiminase pathway leading to the production of putrescine and carbamoyl phosphate from carbamoylputrescine. Carbamoyl phosphate can be converted into ammonium and carbon dioxide with production of ATP through the action of a carbamate kinase (such as ArcC3). The *otcA* gene has been proposed to be renamed to *ptcA*, while the proposed functions of the products of *yrfC* and *yrfD* are an amino acid permease and agmatine deiminase, respectively [31]. A region of approximately 15 kb encompassing these

genes is lacking in *L. lactis* MG1363. This region is not present in the *L. lactis* SK11 draft assembly but could be present in an unsequenced region.

Another gene present in *L. lactis* IL1403 but absent in *L. lactis* MG1363 is *umuC*, a gene involved in UV-inducible repair of mutations in DNA. This gene has previously been described to reside on a plasmid in *Lactococcus*: it is present on pNP40 of *L. lactis* DRC3 [32] and on pIL7 of *L. lactis* IL594, the parent of *L. lactis* IL1403 (Bardowski J. and Koryszewska A., personal communication). Furthermore, it is present in the *L. lactis* SK11 draft assembly. The entire region surrounding *umuC* is absent in *L. lactis* MG1363 suggesting that it was lost as a result of a deletion event. Alternatively, *L. lactis* IL1403 may have acquired the gene through an (plasmid) insertion event. Both strains contain *dinP*, a paralog of *umuC*.

3.4. IS element-derived mutations in *L. lactis*

It is clear now from the genomic information that several of the changes and genetic rearrangements can be explained by the earlier mentioned abundance in *L. lactis* of mobile genetic elements such as prophages and IS elements as well as the presence of conjugative transposons in several strains. The variation in the copy number of the various IS elements and their position in the genomes of related strains of *L. lactis* suggests that at least some of these elements are still active [33].

IS elements can be the cause of several types of mutations and as such could play an important role in bacterial adaptation and evolution [34]. They can disrupt genes, lead to an increase in gene expression when an element carries an outward-directed promoter and inserts upstream of the particular gene, or result in chromosomal rearrangements (insertions, deletions, inversions or duplications) through recombination of different copies of the same element. An analysis of all of the IS elements and their surrounding regions in the chromosome of *L. lactis* IL1403 revealed that two of the 15 IS983 copies and six of the 10 IS981 copies correspond to deletion events between tandem repeats of the respective elements [9]. Thus, in contrast to the situation described above for the NCDO712 family of strains in which IS elements were shown to be involved in a large chromosomal inversion, IS element activity in *L. lactis* IL1403 seems to have led to only relatively small deletions involving proximal copies of IS981 or IS983.

At least one of the IS elements of *L. lactis* IL1403, IS981, was shown to be active under laboratory conditions and caused a number of different chromosomal changes under both normal and stressed growth conditions for prolonged periods of time (approximately 1000 generations in five months for the growing cultures [35]). Southern hybridization revealed that IS904 and

IS983 were not mobile under the conditions employed. The IS981-mediated mutations, two insertions at new positions and five deletions, were all caused by replicative transpositions of the element, indicating that this is the preferred mobility mechanism of IS981. A molecular analysis revealed that in all cases but one these mutations disrupted the function of one or more genes. At the moment it is not clear whether these IS element-induced mutations led to the fitness benefit that resulted in the mutants becoming predominant in the evolved cultures or whether the IS mutations hitched to high frequency along with beneficial mutations elsewhere in the chromosome [35]. IS981 was also shown to be active in an *L. lactis* MG1363 background [36]. In an *ldh* knockout strain grown for a prolonged period of time under anaerobic conditions, growth on glucose and concomitant lactate production were restored after a variable number of generations. This was shown to be the result of transcriptional activation of a silent *ldhB* gene encoding an Ldh enzyme with kinetic parameters different from that of lactate dehydrogenase encoded by the *las* operon. In half of the cases examined, the transcriptional activation was caused by the insertion of a copy of IS981 upstream of *ldh*, at exactly the same position in all instances. By the insertion an IS-derived –35 promoter sequences was fused at the proper spacing to a –10 sequence present in the *ldhB* upstream sequence, creating a functional promoter to drive *ldhB* expression.

4. Global analyses of *L. lactis* – transcriptomics

The availability of the genome sequences of various (micro)organisms is rapidly changing the way in which we look at the organisms. Technologies from the genomics arena allow performing research in a global all-embracing way. DNA array technologies, backed up by rigorous statistics and bioinformatics, enable examining the reaction to modifications in genetic make up of the cell through mutations of any sort, or to changes in the environment of any kind, at the level of the so-called transcriptome: levels of all transcripts in an organism at a given point in time. On the other hand, rapid developments in the field of high-throughput protein analysis allow doing the same at the level of the proteome, the total protein content of a cell at a particular moment in its life cycle.

DNA macroarrays as well as DNA microarrays have recently been devised and used in the analysis of transcription and gene regulation in *L. lactis* and will be discussed here. Several of these studies have been completed and are in the various stages of publication at the time of writing of this review. As they represent the vanguard of research in this particular area of research in *L. lactis* they have been included here.

4.1. Transcriptome analyses in *L. lactis* using partial DNA arrays

Pioneering work on metabolic flux control using a limited number of genes in an extended Northern dot blot (“DNA macroarray”) format was performed by Even et al. [37–39] and Fontaine et al. [40]. Due to the restricted number of genes analyzed these studies do not qualify as genuine post-genomic approaches. Nevertheless, they have uncovered a number of important considerations with respect to mRNA stability that warrant more extended DNA (micro)array studies. Nylon membranes carrying 500-bp amplicons of 15 genes from the glycolysis and mixed-acid-fermentation pathways of *L. lactis* IL1403 were used to characterize sugar catabolism and the response to acid of *L. lactis* IL1403 [37–40]. An alkaline phosphatase-based enzyme assay using phosphorfluorimaging was employed to measure the abundance of a specific transcript in the total RNA pool. Cellular concentrations of mRNAs were derived from correcting for the cellular RNA concentration. In the first study, the metabolic characteristics of *L. lactis* IL1403 growing exponentially in two different chemically defined media with glucose or galactose were thoroughly examined by following growth kinetics, intracellular metabolites, enzyme activities and gene expression [37]. Homolactic fermentation was observed in all cases due to either the low concentration of pyruvate formate lyase or of alcohol dehydrogenase. The observed sugar consumption rates were maximal rates suggesting that the catabolic flux could not be increased, a problem that may be at the level of transport of the sugars. The concentration of total RNA was generally higher in glucose- than in galactose-grown cells. When the data were corrected for this phenomenon, it was shown that several genes of glycolytic enzymes were more highly expressed during growth of *L. lactis* IL1403 on glucose. Conversely, transcripts of genes of the mixed-acid pathway were more abundant during growth on galactose. Comparing the cellular transcript concentration to the rate of enzyme synthesis, an average increase in translational efficiency of 3-fold was observed of glucose- over galactose-grown cells, suggesting that mRNA translation is regulated by a sugar-dependent mechanism [37].

In a study of the response of *L. lactis* to acid stress *L. lactis* MG1363 was grown anaerobically with a large excess of glucose [38]. The culture remained homolactic throughout. The decrease of the culture pH as a consequence of glucose metabolism led to cytoplasmic acidification and to inhibition of the activities of enzymes of the glycolytic and mixed-acid pathways. The consequent decrease in catabolic flux and energy production, together with the extra energy used to counteract the acidification of the cytoplasm, resulted in a diminished biomass production. The specific growth rate decreased

until the culture ultimately stopped growing, although remaining glucose was slowly metabolized in the post-acidification phase. Gene expression data on five different time points of the culture revealed that the transcripts concentration of most of the metabolic genes decreased significantly. The stability of the transcripts, measured after the addition of the transcription initiation inhibitor rifampicin, of a number of genes (*glk*, *gap*, *pgk*, *enoA*, *ldh*, *pfl* and *adhE*) was markedly (2- to 10-fold) increased during the post-acidification phase. A change in mRNA turnover was not apparent for the other genes examined. Transcription at reduced rates, together with increased transcript stability, during the post-acidification phase results in *L. lactis* maintaining a pool of mRNA all set for translation. This could represent a survival mechanism: it would allow the cells to maintain a minimal level of catabolic activity under energy limitation and to rapidly recover and resume growth once the limitation (stress) is removed [38]. It would be very interesting to see the extent of this response by using full genome DNA microarrays.

By growing *L. lactis* MG1363 as pH-stat cultures (pH 6.6, 5.5, 5.2 and 4.7) on glucose medium, cellular responses to pH were uncoupled from changes in the growth rate arising from medium acidification in batch cultures [39]. A significant amount of glucose accumulated in the medium at pH 4.7; above this pH the cultures were carbon-limited, with a progressive shift towards homolactic fermentation with decreasing steady-state pH. The growth-limiting factor at pH 4.7 was lactate, which was most inhibitory under acidic conditions of growth. Acceleration of glycolysis, one of the main cellular responses to acidification, was regulated by an increase in the concentrations of the glycolytic enzymes and the specific modulation of enzyme activities. The main factor controlling pathway flux under acidic conditions was shown to be metabolic regulation: it allows the flux to be maintained even though enzyme activities are decreased due to the low intracellular pH. The enzyme concentrations increased while the mRNA concentrations remained constant or reduced slightly under acidic conditions. Translation efficiencies of a number of metabolic genes were calculated and shown to increase at low pH. This is presumably accompanied by an increase in ribosomal activity and, together with the observed reduction in the concentration of total RNA at the lower pH values it suggests that the translation apparatus of *L. lactis* is optimized under acidic conditions by the optimization of active ribosomes.

The studies presented above show that care has to be taken, especially when interpreting full-genome transcript profiles, as the level of mRNA measured not only depends on the level of transcription but also on potentially differential mRNA stability.

A partial amplicon-based DNA microarray of *L. lactis* MG1363 was devised that contained genes involved

in nucleotide biosynthesis (50), energy metabolism (59), environmental stress response (24) and other areas of cellular metabolism (79) (Kilstrup, M., Nielsen, A.K., Pulka-Amin, M.D. and Rasmussen, M.D., unpublished). These authors examined the effects of nucleotide deficiency in *L. lactis* MG1363, imposed by inhibition of GMP synthase activity by decoynine (GTP depletion) or by exhausting the supply of the purine base hypoxanthine in a *purD* mutant (GTP and ATP depletion). In the latter case, these two precursors of RNA synthesis will be severely limited and total RNA synthesis is expected to decline and ultimately stop, invalidating internal normalization procedures. Indeed, cellular rRNA levels were shown to vary slightly: they increased during purine depletion in the *purD* strain and decreased after decoynine addition to MG1363. Details of the results of this work are presented elsewhere in this issue (Martinussen et al. [122]). The array was also used to develop a normalization method that takes into account the fact that the level of rRNA, which represents approximately 85% of the total cellular RNA, may vary more than 3-fold in response to growth conditions [41–43]. The method measures rRNA content at the single cell level using a Cy3-labeled probe specific for 16S rRNA from *L. lactis* ssp. *cremoris* in quantitative fluorescence in situ hybridization, coupled to automated image analysis and signal quantification. The method should be generally applicable in circumstances where internal normalization (e.g., when using a partial DNA microarray to examine conditions in which expression of many or most of the genes represented on the array is expected to change) is impossible (Kilstrup, M., Nielsen, A.K., Pulka-Amin, M.D. and Rasmussen, M.D., unpublished).

A partial DNA macroarray of 22- to 24-mer oligonucleotides represented 375 genes of *L. lactis* IL1403 involved in protein degradation, carbohydrate-, fatty acid-, nucleic acid-, or amino acid metabolism [44]. Three stress conditions of 30 min each were applied to the cells in GM17 medium: heat (42 °C), acid (pH 5.5 with lactic acid) or osmotic stress (4% NaCl). There was good agreement of stress-induced gene expression changes observed by the DNA macroarrays and what was already known from literature, adding a level of confidence to the new responses observed with the DNA macroarrays. Heat shock led to changes in expression of 64 genes involved in the nucleotide salvage pathway, methionine biosynthesis, glutamate uptake and biosynthesis, arginine catabolism and betaine uptake [44]. Expression of a number of putative ABC amino acid transporters was increased while two peptide transporters were strongly repressed. Expression of 50 genes was altered during the 30-min lactic acid treatment. The genes of the two-peptide transporters, and the *pepX* gene specifying the peptidase PepX, were repressed. Expression of the β -glucoside-specific phosphotransfer-

ase system specified by *yedEF* was also strongly (over 100-fold) repressed. Other repressed genes were those of glutamine biosynthesis and transport, *glnA* and *glnPQ*, respectively. Interestingly, the latter genes had already been implicated in lactic acid resistance in lactococci [45]. Among the induced genes were those for choline transport (*choQS*), tryptophan biosynthesis and citrate and malate fermentation (*citE*, *citF*, *mleS* and *ldh*). Thirty-four genes were induced by osmotic stress while another 34 were repressed under these conditions. The genes *busAA-AB*, specifying a high-affinity betaine uptake system [46], were induced approximately 50-fold. Other induced genes were those for nucleotide salvage, biosynthesis of the amino acids lysine and glutamate, peptidoglycan biosynthesis and, as expected from earlier results, DnaK. Genes of the arginine deiminase pathway (see below), fatty acid biosynthesis and those involved in citrate and malate fermentation were repressed to various extents. Again, as was the case for the other two stress conditions, the genes for a number of proteolytic enzymes, and *yedEF*, were repressed during osmotic stress.

4.2. Nitrogen metabolism

4.2.1. Proteolysis – CodY

L. lactis is an amino acid auxotroph: most strains require an exogenous supply of various amino acids, either in the free form or contained in (oligo)peptides. An elaborate proteolytic system consisting of an extracellular proteinase, intracellular peptidases and membrane-located peptide uptake systems allows the organisms to degrade milk proteins and internalize the required amino acids and peptides [47]. Due to its importance in flavour development in cheese manufacture, proteolysis has been studied in detail for many years and the genes of most enzymes and proteins involved had already been cloned, sequenced and analyzed before whole genome sequencing took off. The available genome sequences do allow to look at the proteolytic genes in their genetic context more thoroughly and to understand the genetic instability of certain traits (e.g., that of *opp-pepO*, see above). The emphasis of research in this field has changed towards understanding the intricacies of proteolysis regulation. A central regulator, CodY, has been identified as the main player in this respect [48–50]. In nitrogen-rich media expression of a number of proteolytic genes is repressed, while expression is relieved under nitrogen limitation. CodY acts as a transcriptional repressor of the expression of *prtMP*, *opp*, *pepN*, *pepC*, *araT* and *bcaT* [49,51–53]. The signal affecting the strength of repression by CodY lies in the intracellular pool of branched-chain amino acids (BCAA), which act as cofactors stimulating CodY binding to its operator site upstream of target genes

[50,51]. Direct in vitro binding of multiple molecules of CodY to regions upstream of several of its target genes/operons has been shown to occur and to depend on the presence of BCAA in the assay buffer, both by electrophoretic mobility shift assays and by DNase I footprinting experiments [51].

These genetic studies have recently been complemented with a genome-wide analysis aiming at elucidation of the CodY regulon, using full genome DNA microarrays of *L. lactis* [54,55]. To this end, the transcriptomes of an *L. lactis* MG1363 *codY* mutant, carrying a 432-bp internal deletion in *codY*, and its isogenic parent were compared after growth of both strains to exponential phase in a rich medium with excess nitrogen, i.e., in which CodY exerts strong repressing activity [51]. The labeled cDNAs were hybridized to the DNA microarrays carrying 2145 amplicons of genes of *L. lactis* IL1403. The expression of over 30 genes was significantly altered between the strains under these conditions. Although the majority of the de-repressed genes are involved in amino acid transport and metabolism, the expression of several genes belonging to other functional categories was also changed in the mutant, emphasizing the pleiotropic role of CodY in *L. lactis*. One of the strongest de-repressed transcriptional units was *opt*, which is over 90% similar at the nucleotide level to *dpp* of MG1363, and was previously shown to be repressed in peptide-rich media and especially in media containing peptides with BCAA residues [49,56]. The *leu-ilv-ald* gene cluster was significantly de-repressed (over five times) in the *codY* mutant, as was the amino acid transferase gene *bcaT*. The gene for the putative asparagine synthetase *AsnB* and those for *GltDB*, presumed to catalyze the same reaction, were elevated four and six times, respectively, upon functional removal of *codY*. In addition to the peptide transporter genes, the gene for a putative cationic amino acid transferase was de-repressed five times in MG1363*codY*. Interestingly, expression of genes of the Krebs oxidative cycle (*gltA*, *icd*, *citB*, encoding citrate synthase, isocitrate dehydrogenase and aconitase, respectively) is also de-repressed in the *codY* strain (den Hengst, C.D., van Hijum, S.A.F.T., Gerts, J.M.W., Nauta, A., Kok, J. and Kuipers, O.P., unpublished). The authenticity of a number of the identified new CodY targets was verified by *lacZ* fusion studies and by electrophoretic mobility shift assays using purified CodY protein.

An in silico analysis for common sequence motifs in the upstream regions of a number of known and new CodY target genes, as identified by the DNA microarray study, enabled the definition of a CodY binding box (den Hengst, C.D., van Hijum, S.A.F.T., Gerts, J.M.W., Nauta, A., Kok, J. and Kuipers, O.P., unpublished). A homologue of *L. lactis* CodY, CodY of *B. subtilis*, is a nutritional repressor of the dipeptide permease operon (*dpp*), and of genes involved in amino acid

metabolism, carbon and energy metabolism, antibiotic production, competence development and motility in *B. subtilis* [57–67]. Like *L. lactis* CodY, affinity of *B. subtilis* CodY for its targets is stimulated by interaction with BCAAs but also through a BCAA-independent interaction with GTP [57,62,68,69]. Thus, *B. subtilis* CodY is thought to sense both nitrogen availability and the energy state of the cell. Using the lactococcal CodY box sequence derivatives of this box were also shown to be present in the upstream sequences of a number of known targets of *B. subtilis* CodY.

4.2.2. *GlnR*-mediated regulation

Glutamine and glutamate are the nitrogen donors for most nitrogenous molecules in the bacterial cell. Both amino acids are interconverted via glutamine synthetase (GS) and glutamate synthase GOGAT: glutamate and ammonium are converted to glutamine, at the expense of one ATP, by GS, while GOGAT converts glutamine and α -ketoglutarate into two molecules of glutamate (see Fig. 3). This circular pathway is at the intersection between carbon and nitrogen metabolism and is subject to complex regulation that is executed in fundamentally different ways in the various bacterial species in which it has been studied [70–74]. It is outside the scope of this review to go into the intricacies of these different regulatory mechanisms and we will focus only on a comparison of GlnR-mediated control in *B. subtilis* and *L. lactis*.

Nitrogen control in *B. subtilis* is exerted via the three transcriptional regulators TnrA, GlnR and the previously mentioned CodY [60]. The former two regulators are members of the MerR family of transcriptional regulators and possess highly similar N-terminal DNA binding domains with which they recognize the same transcriptional operators. Due to their differing C-terminal sensory domains they do so, however, in response to different cellular signals. TnrA acts as a repressor or activator when nitrogen is limiting, while GlnR is a repressor of, among others, *glnRA* and *tnrA* under conditions of nitrogen excess. GS was shown to sequester TnrA, thus preventing the latter from binding to its operators, in the presence of a number of feed-back inhibitors of GS activity.

The genomes of *L. lactis* strains MG1363 and IL1403 encode GlnR but not TnrA. The entire *glnR* gene was removed from the chromosome of *L. lactis* MG1363 and the transcriptome of the resulting strain, MG Δ *glnR*, was compared with that of its parent [75]. The strains were grown in chemically defined medium with limiting (0.1%) or excess (2%) casitone as the nitrogen source. Expression of *glnA*, the gene encoding GS and which is located downstream of *glnR* in what appears to be a *glnRA* operon, is de-repressed in *L. lactis* MG Δ *glnR* in both media. The *amtB**glnK* genes, encoding a putative ammonium transporter and a nitrogen regulator PII

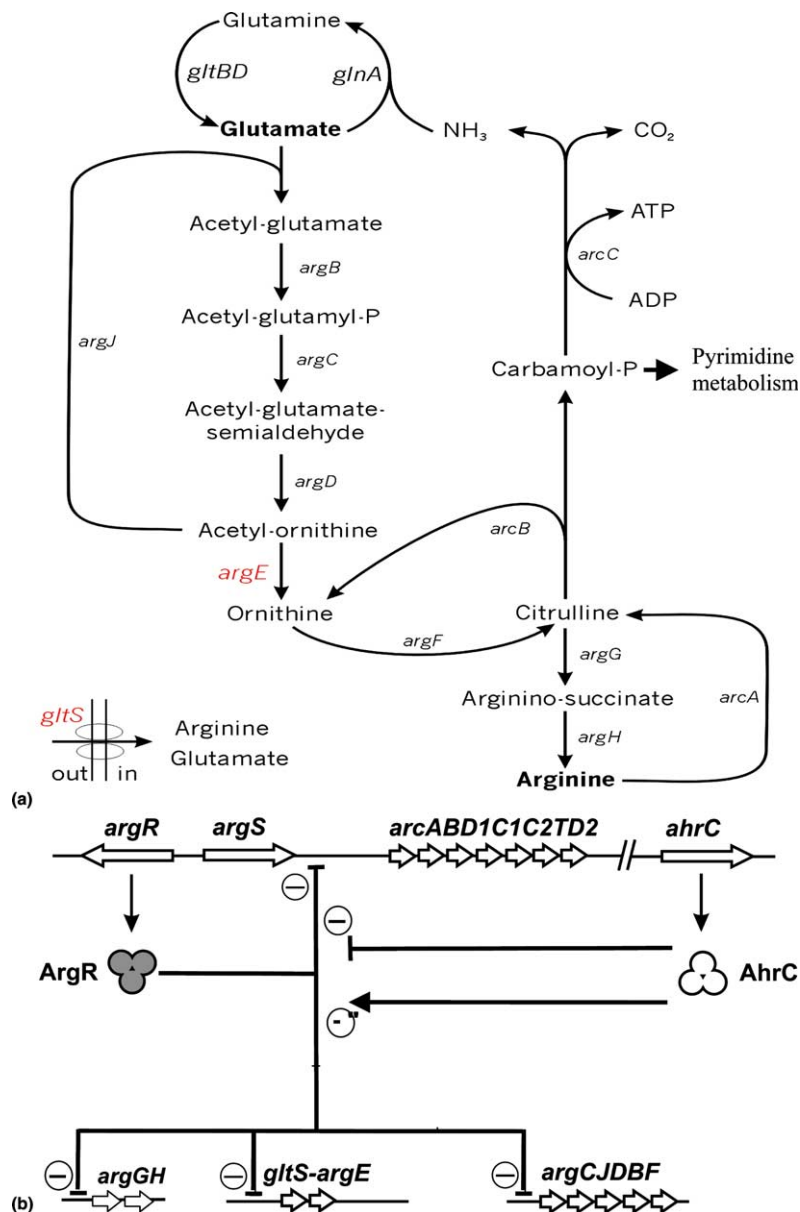


Fig. 3. (a) Schematic representation of arginine metabolism in *L. lactis*. Genes with encoded enzymes: *argB*, *N*-acetylglutamate 5-phosphotransferase; *argC*, *N*-acetylglutamate 5-semialdehyde dehydrogenase; *argD*, *N*2-acetylornithine 5-aminotransferase; *argJ*, ornithine acetyltransferase; *argE*, acetylornithine acetyltransferase; *argF*, ornithine carbamoyltransferase; *argG*, argininosuccinate synthetase; *argH*, argininosuccinase; *arcA*, arginine deiminase; *arcB*, ornithine carbamoyltransferase; *arcC*, carbamate kinase; *gltS*, arginine or glutamate transporter; *gltDB*, glutamate-oxoglutarate amidotransferase; *glnA*, glutamine synthetase. (b) Model of the possible regulatory mechanism exerted by ArgR and AhrC of *L. lactis*. Positive and negative regulation is indicated by plus and minus signs, respectively. Adapted from [81].

protein, respectively, are de-repressed in low-casitone medium but not affected in high-casitone medium. The *glnPQ* genes encoding a possible glutamine ABC transporter and -binding protein were similarly but less strongly affected by the *glnR* mutation. The *glnR* and *amtBglNk* upstream regions contained 18/19-bp sequences with similarity to the 17-bp GlnR box of *B. subtilis*. Subsequently, this motif was also identified in the start of *glnP*. A multi-copy (plasmid-located) transcriptional *lacZ* fusion to the upstream region of *amtBglNk* confirmed the GlnR dependency of expression from this

region at a low (0.1%) casitone concentration in the medium. Regulation was disrupted when half of the GlnR box upstream of the -35 region in the *amtBglNk* promoter was deleted [75]. Chromosomal transcriptional *lacZ* fusions to *amtBglNk*, *glnRA*, and *glnPQ* in *L. lactis* MG1363 and *L. lactis* MGΔ*glnR* were used to study the regulatory effects of glutamine, glutamate and ammonium in a single-copy situation. These studies revealed that regulation of *glnRA* and *amtBglNk* is quite different. Repression of *glnRA* in response to glutamine and ammonium was GlnR dependent. This was also the

case for glutamine-mediated repression of *amtBglN*K expression. However, ammonium still had a repressive effect on *amtBglN*K, leading to the conclusion that an additional, ammonium-responsive factor is involved in regulation of this operon.

4.2.3. Arginine metabolism – ArgR/AhrC

Arginine metabolism in *L. lactis* consists of a biosynthetic route producing arginine from glutamate and a catabolic or arginine deiminase pathway, in which arginine is degraded to carbamoyl phosphate, ammonium, CO₂ and ATP (see Fig. 3(a)). The ATP is an important alternative source of energy, while carbamoyl phosphate is a precursor in the de novo synthesis of pyrimidines. The ammonium counteracts the lowering of the pH as a consequence of sugar fermentation and the arginine deiminase pathway may be involved in acid stress resistance [76–80]. The arginine biosynthesis genes are *argCJDBF*, *gltS-argE* and *argGH*, while the three steps of arginine breakdown are encoded by *arcA* (for arginine deiminase), *arcB* (for ornithine carbamoyltransferase) and *arcC1C2* (for carbamate kinase). The latter genes are part of the cluster *arcABD1C1C2TD2* in which *arcD1* and *arcD2* specify putative arginine–ornithine (1:1) antiporter proteins and *arcT* encodes an aminotransferase.

Two arginine regulators, ArgR and AhrC, are operative in *L. lactis* [81]. They are 32% identical and homologous to the arginine regulator ArgR from *E. coli* and *B. stearothermophilus* and to AhrC from *B. subtilis*. A *lacZ* fusion to the upstream region of the *argCJDBF* genes was used to study arginine biosynthesis gene regulation as a function of arginine availability and the role of the two regulators therein. Expression from the *argC* promoter was arginine-dependent in *L. lactis* MG1363 but was de-repressed in either of the two single regulator mutants (*L. lactis* MGΔ*argR* and *L. lactis* MGΔ*ahrC*) and in the double mutant *L. lactis* MGΔ*argR*Δ*ahrC*. Apparently, the two regulators cannot complement each other. Using a *lacZ* fusion to the upstream region of *arcA*, the first gene of the catabolic *arc* gene cluster, it was shown that expression from this region was increased in the Δ*argR* strain, while it was strongly down-regulated in the Δ*ahrC* strain. Thus, both regulators seem to have opposite roles in arginine catabolic gene regulation. As an *argRahrC* double mutant had the same phenotype as the Δ*argR* strain, i.e., it was de-repressed for *arc* gene expression, AhrC is apparently not essential for activation of the arginine deiminase pathway [81]. Transcription profiles were obtained for the three regulator mutants, MGΔ*argR*, MGΔ*ahrC* and MGΔ*argR*Δ*ahrC*, and compared to those of the parent strain *L. lactis* MG1363 [75]. As the largest effects of the mutations on gene regulation were observed with high amounts of arginine [81], RNA was isolated from cells growing in chemically defined medium containing

10 mM arginine. The results indicate that ArgR and AhrC are dedicated regulators of arginine metabolism as high expression ratios (mutant over wildtype) were only observed for genes known to be involved in this metabolic pathway (but see below). The arginine biosynthesis genes *argCJDBF*, *gltS-argE* and *argGH* were de-repressed in all three mutants. The *argR* deletion had no significant effect on the arginine catabolic genes while Δ*ahrC* led to down-regulation of *arcA* and *arcD1C1C2*. In the double mutant MGΔ*argR*Δ*ahrC* the entire *arcABD1C1C2TD2* cluster was significantly de-repressed. Thus, AhrC is involved in activation of the arginine deiminase pathway but the results obtained with the double mutant show that ArgR is also somehow involved.

Regulation of arginine metabolism has been studied extensively in *E. coli* and *B. subtilis*. The de novo biosynthesis of arginine is repressed by the transcriptional regulators ArgR and AhrC in *E. coli* and *B. subtilis*. In both bacterial species regulation occurs through binding of the respective regulator to an 18-bp imperfect inverted repeat structure that overlaps the core promoters of the regulated genes. These operators are called ARG boxes [82–85]. As the consensus sequences of these ARG boxes from *E. coli* and *B. subtilis* were too divergent to predict the sequence of a lactococcal ARG box, the data of the DNA microarray experiments were used to find such operators in *L. lactis* [75]. Over-represented sequences in the regions upstream of the first genes of the ArgR or AhrC regulated operons were searched for with the Gibbs Motif Sampler [86]. ArgR- and AhrC-box motifs were found that contained a weak palindromic structure as well as residues conserved in known ARG boxes. Subsequently, these sequences were used to search all promoter regions of *L. lactis* MG1363 using matrices built on either consensus. Eight out of the 10 boxes with highest similarity to the ArgR regulon motif were present upstream of genes of the arginine biosynthesis pathway, while the same was the case for four of the seven highest scoring “AhrC motifs”. Interestingly, the *arcA* promoter region contained a relatively weak AhrC-motif consensus, but it does contain several ARG operator half-sites with high similarity to the AhrC operator of *B. subtilis*.

The *argR* and *ahrC* mutations significantly affected the expression of only a small number of other gene clusters, namely those involved in pyrimidine de novo biosynthesis (*pyrRPBcarA*, *pyrEC* and *pyrKDbF*) and a cluster of genes with putative functions only (*ybdA-rnpA-ybdCD*). In the *argR* and the regulator double mutant, expression of pyrimidine biosynthesis genes was increased relative to wildtype, while no effect was observed in the Δ*ahrC* strain. The DNA microarray results were confirmed by analyses of PyrE (orotate phosphoribosyltransferase) and PyrF (orotidine-phosphate decarboxylase) enzymatic activities in the various strains [75]. Carbamoylphosphate is the common metabolite in

arginine and pyrimidine metabolism: it is used during arginine biosynthesis in the production of citrulline from ornithine (via ArgF) and is a precursor in pyrimidine biosynthesis (see Fig. 3(a)). Carbamoylphosphate is produced from citrulline during arginine degradation (via ArcB) or from ammonium, CO₂ and ATP by carbamoylphosphate synthase (specified by *carAB*). The *car* and *pyr* genes are regulated by PyrR in response to pyrimidines, not to arginine [87]. The citrulline levels were relatively high in the Δ *argR* and the Δ *argR* Δ *ahrC* strain and low in Δ *ahrC*, which correlates well with the expression seen for both biosynthetic pathways and less so with that of arginine degradation. Most likely, the increased flux towards arginine in the Δ *argR* and Δ *argR* Δ *ahrC* strains leads to carbamoylphosphate and pyrimidine depletion, which is counteracted, apparently, by the induction of the pyrimidine biosynthetic genes including *carAB* [75].

Interaction between both purified His₆-tagged arginine regulators was studied by DNase footprinting and gel retardation assays [88]. AhrC alone did not bind to DNA while ArgR bound to the promoters of the biosynthetic *argCJDBF* and the catabolic *arc arcABD1C1C2TD2* gene cluster, independent of the presence or absence of arginine. Arginine-dependent in vitro DNA binding was obtained when both purified regulators were mixed. In that case, addition of arginine resulted in increased binding of ArgR/AhrC to the biosynthetic *argC* promoter but to a decreased binding to the catabolic *arcA* promoter. The ArgR/AhrC complex protected 2 ARG box operators in the upstream region of P_{*argC*} in a DNase footprint. In the absence of AhrC, ArgR protected several ARG box half-sites in promoter region of *arcA*.

A model was proposed to describe the functions of ArgR and AhrC in arginine-mediated transcriptional regulation in *L. lactis* (Fig. 3(b)). In the absence of arginine ArgR is thought to bind only to the catabolic *arcA* promoter, preventing arginine degradation via ADI. The three arginine biosynthetic operons are left unrepressed, resulting in arginine production. In the presence of arginine ArgR and AhrC form a complex with high affinity for ARG boxes, leading to de-repression of the *arc* genes and utilization of arginine via ADI and repression of arginine biosynthesis [81,88].

4.3. Identification of a novel glucose-PTS system in *L. lactis* by transcriptome analysis

As soon as genetic engineering protocols became available for *L. lactis* efforts have been undertaken to manipulate certain fermentation pathways to either (better) understand the routes or, on the basis of this knowledge, improve the organism for use in industrial applications. Galactose and lactose are not always desirable in food products. Lactose intolerance is a well-known phenome-

non while the effect of consumption of galactose on cataract occurrence in certain individuals, especially in combination with alcohol intake, has been documented [89]. Natural sweetening of food products by fermenting organisms can be a desirable trait in some applications. The combination of considerations mentioned above led to an attempt to engineer *L. lactis* MG1363, growing on lactose, into a glucose-producing cell factory with concomitant lowering of lactose and galactose levels in the medium. A strategy was designed to remove all glucose-phosphoenolpyruvate transport system (PTS) import activities. The glucokinase gene was also deleted to prevent glucose, entering the cell through non-PTS, from being phosphorylated and further processed via glycolysis (Pool, W.A., Neves, A.R., Kok, J., Santos, H. and Kuipers, O.P., unpublished). EII^{man/glc} (*ptnABCD*) is able to import glucose [90] but a Δ *glk* Δ *ptnABCD* double mutant of *L. lactis* MG1363 was found to still be able to grow on glucose (Pool, W.A., Neves, A.R., Kok, J., Santos, H. and Kuipers, O.P., unpublished). This pointed to the involvement of another PTS(s) in glucose uptake. As several candidate PTSs might be involved in glucose transport, DNA-microarrays were employed to compare the transcriptomes of *L. lactis* strains NZ9000 (a *nisRK*-containing derivative of *L. lactis* MG1363) and NZ9000 Δ *glk* Δ *ptnABCD* grown in CDM with 1% (w/v) glucose. This analysis unambiguously pointed to the putative PTS EII^{cel} (encoded by *ptcBAC*) as being the only PTS system significantly (>5 times) up-regulated in the double knock-out strain. Cellobiose is a disaccharide made up of two glucose units and, apparently, EII^{cel} of *L. lactis* can transport both sugars. Based on these results a triple knock-out strain (*L. lactis* NZ9000 Δ *glk* Δ *ptnABCD* Δ *ptcBAC*) was selected on galactose. Indeed, this strain was not able to grow on glucose. Subsequently, the lactose plasmid pMG820 [91], carrying the genes for the lactose-PTS and the tagatose-6-phosphate pathway, was introduced in the triple mutant. This strain was shown to grow somewhat slower on lactose but it efficiently fermented the galactose moiety of lactose, while excreting glucose at levels equimolar to the lactose internalized (Fig. 4). No galactose was detectable in the growth medium as opposed to the wild-type strain carrying pMG820, which produced a low amount of galactose. Notably, the engineered strain was able to grow in skim milk and produced up to 22 mM of glucose in the medium. This study shows that even when a full genome sequence is available, missing links in a metabolic engineering strategy are not trivial to find. Transcriptome analysis is then a suitable and fast way to identify factors for further directed engineering.

4.4. Fructose metabolism in *L. lactis*

A great deal of attention on sugar metabolism in lactococci has gone, obviously, in the fermentation of the milk sugar lactose, and in the breakdown by these

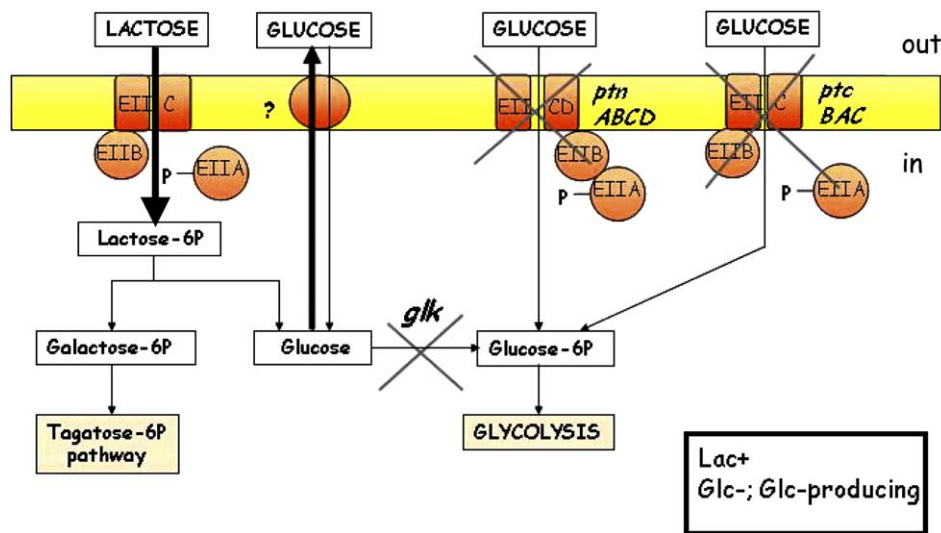


Fig. 4. Routes in *L. lactis* MG1363 for utilization of lactose and glucose. The depicted *L. lactis* MG1363 derivative, *L. lactis* NZ9000AglkAptnABCDptcBAC (pMG820) (Pool, W.A., Neves, A.R., Kok, J., Santos, H. and Kuipers, O.P., unpublished), can grow on lactose (Lac⁺) as the genes for lactose-PTS and the tagatose-6-phosphate pathway are present on the plasmid pMG820. Due to the mutations (gray crosses) in the strain it cannot grow on glucose (Glc⁻). It efficiently ferments the galactose moiety of lactose and excretes glucose at levels equimolar to the lactose supplied (thick arrows). Question mark: the number and identity of glucose permeases in *L. lactis* is largely unknown. For details on gene function: see text.

organisms of glucose. Much less is known about the utilization of fructose, a sugar predominantly associated with plants. Many species of bacteria can use fructose as a carbon and energy source. Fructose utilizing ability allows for enhanced survival of, e.g., oral streptococci in the oral cavity while it has also been implicated in bio-film formation of *S. gordonii* in that environment [92]. Although permease systems have been described, fructose utilization most often involves specific PTSs consisting of an EII^{Fru} enzyme and 1-phosphofructokinase, the genes of which are encoded by genetically very similar *fru* operons in different bacterial genera. The two genes are commonly accompanied by a gene specifying a transcriptional regulator of the DeoR family. Surprisingly, this regulator (designated FruR) seems to have opposite roles in *Spiroplasma citri*, a plant pathogen, and *S. gordonii*, a commensal of the oral cavity [92,93]. While in both cases fructose enhances transcription of the respective operons, FruR is an activator in *S. citri* and acts as a repressor in *S. gordonii*.

Recently, fructose utilization has been studied in considerable detail in *L. lactis* IL1403 using, among others, global transcript analysis [94]. In the course of that work, the authors renamed the genes involved on the basis of the obtained results: *fruR*, *fruC* (formerly *lacR* and *lacC*, respectively) [21] and *fruA*. Single mutants in *fruC* and *fruA* have established that the *fruRCA* specifies the main fructose utilization pathway in *L. lactis* IL1403. However, functional analysis of these mutants also indicated that *fruRCA* is not the only fructose utilization pathway encoded in this strain. Expression from the *fru* promoter was strongly de-repressed on glucose in a

fruR mutant, indicating that FruR acts as a repressor in *L. lactis*, while the *fru* operon is under catabolite repression by CcpA. As non-polar mutations in *fruA* and/or *fruC* lead to constitutive low expression of the *fru* operon, even in the presence of fructose, another level of control is exerted at the metabolic level, most likely through fructose-1-phosphate. To define the entire FruR regulon the transcriptomes of *fruR* and *fruC* single mutants of *L. lactis* IL1403 grown in chemically defined medium with trehalose were compared with that of the parent strain [94]. In both mutants the only genes differentially expressed were those of the *fru* operon. Apparently, FruR is a dedicated regulator and involved in fructose utilization only. Examining the upstream regions of *fru* operons of gram-positive bacteria with orthologous FruR regulators pinpointed four adjacent 10 bp direct repeats that possibly constitute a FruR binding site. Indeed, two mutated versions of the region led to de-repression of *fru* transcription in *L. lactis*.

4.5. Transcriptome analysis reveals nisin-resistance mechanisms of *L. lactis*

L. lactis can be readily made resistant to nisin under laboratory conditions. Resistance development is a reversible process: when an *L. lactis* Nis^R strain, which is 75-fold more resistant to nisin than its parent, is grown overnight without nisin the resistance drops to almost the original level. Thus, nisin resistance seems to be a reversible property rather than to be caused by mutations in *L. lactis* and, as for that matter, also in a number of other gram-positive bacteria (Kramer,

N.E., personal communication). To investigate which mechanism(s) *L. lactis* employs to become temporarily resistant to nisin, DNA microarrays of *L. lactis* IL1403 were employed to compare the transcriptomes of *L. lactis* IL1403 and its isogenic derivative *L. lactis* IL1403 Nis^R. During the experiments the latter was grown in the presence of 3 mg nisin/l. The analysis revealed the significant up/down regulation of 95 genes, comprising 4.4% of all annotated genes of IL1403 [95]. The main functional categories to which these genes belong are (i) cell wall synthesis, (ii) central and energy metabolism, (iii) phospholipid- and fatty acid metabolism, (iv) gene regulation, (v) transport, (vi) stress, and (vii) miscellaneous or unknown functions. Genes directly involved in lipid II biosynthesis were not differentially expressed: lipid II is the docking molecule that enables nisin to form lethal pores in the cytoplasmic membranes of sensitive cells at nanomolar concentrations [96]. Lipid II is indispensable for cell wall biosynthesis and the pyrophosphate group recognized by nisin is synthesized so early in the biosynthesis route that the cell cannot afford to change the lipid II moiety without severely suffering from the consequences. This notion is also supported by previous findings showing no direct involvement of lipid II levels in acquired nisin resistance [97].

Some of the genes that were most clearly up-regulated in the transcriptome analysis were those of the *dlt* and *gal* operons. Both operons are involved in the substitutions that take place in lipoteichoic acid (LTA) in the cell wall. D-alanylation of LTA, specified by *dlt*, increases the positive charge in the cell wall, thereby probably hampering access of nisin to the cytoplasmic membrane [95]. Strains carrying knockout mutations in these operons displayed a 4- to 6-fold increased sensitivity to nisin, demonstrating the relative importance of these genes. Also the penicillin binding protein encoded by *pbp2A* was found to be up-regulated. This is in accordance with the finding in a previous study in which *Listeria monocytogenes* Nis^R was shown to also overexpress a penicillium-binding protein [98]. Another class of genes that was de-repressed upon nisin resistance development was that specifying ABC transporters. Included in this group are YsaBC (orthologs of MbrBA of *S. mutans*, involved in bacitracin resistance), and YneGH, a transporter putatively involved in arsenic resistance. Interestingly, orthologs of bacitracin resistance genes were also up-regulated in nisin resistant *B. cereus* and *S. pneumoniae*, suggesting an important role of the encoded transport system in generating resistance against antimicrobial peptides (Kramer, N.E., personal communication). Genes involved in nitrogen metabolism were also up-regulated, in particular the arginine catabolic pathway. An increase in the activity of the latter pathway leads to elevated production of ammonia by the cells, which could lead to elevation of the pH in the cell wall. If this is the case it would increase

the affinity of the cell wall for nisin and, thus, hamper the bacteriocin from reaching the cytoplasmic membrane. Knocking out AhrC, one of the regulators of the catabolic *arc* genes (see above) yielded a strain that was, indeed, more sensitive to nisin. Membrane biosynthesis was also found to be affected in the *L. lactis* Nis^R strain, albeit probably only moderately so: the *fab*-operon, responsible for the saturation of fatty acids of phospholipids, was down-regulated, suggesting that more unsaturated acyl chains are built into the membranes of *L. lactis* Nis^R.

Overall, a picture emerges of nisin resistance in *L. lactis*, and probably also in other gram-positive bacteria, as a multi-factorial trait, with several mechanisms operating at the same time. Most of these mechanisms appear to be transient or based on adaptation, since resistance is rapidly lost when resistant strains are grown in a medium without nisin. The identified genes could well be used to engineer strains with a constant high and irreversible level of nisin resistance, or the reverse, to engineer strains with a higher susceptibility to nisin. Indeed, a strain of *L. lactis* overexpressing the *yneGH* genes from a plasmid was approximately 10-fold less sensitive to nisin than the parental strain [95].

4.6. *L. lactis* and food safety

Genetic engineering techniques for *L. lactis* allow modifying the organism for entirely novel (food) applications. Especially, the use of *L. lactis* for medical purposes poses great potential, e.g., for the in situ production of medically relevant biomolecules [3]. *L. lactis* itself being a GRAS (generally regarded as safe) organism, the evaluation of safety aspects of such improved or novel strains that are to be used in possible (future) applications is an important issue. The legislation around new strains and the differences as to how these strains were obtained should determine their acceptability is a matter that is discussed to greater extent elsewhere in this issue.

Strains of LAB with improved properties obtained by ultraviolet (UV) radiation or chemical mutagenesis are widely accepted for use in food and feed [99]. The application of genetically modified LAB obtained by modern techniques like site-directed mutagenesis, however, is highly controversial. With the advent of genomics technology, new methods become available to assess to what extent the improved or new strains are altered relative to their parents. Such data could play a role in a subsequent discussion on safety or acceptability in food, feed or other (medical) applications [99]. Alteration of a genetic locus by a targeted approach such as site-directed mutagenesis is expected to result in fewer changes in gene expression and protein production levels than when the same locus were modified by random techniques such as chemical- or UV treatment. In an EU-funded

project (“Express Fingerprints”; QLK3-2001-01473) the side effects of these three different mutation techniques (UV, chemical treatment and site-directed mutagenesis) were compared for the mutagenesis of three traits in *L. lactis*. The three types of mutants in each trait were compared to the isogenic parent strain to obtain an estimate of the number of genes that were affected by each mutational technique. Global gene expression and protein production patterns in the mutants and the parent strain were determined by DNA microarrays and two-dimensional (2D) gel electrophoresis, respectively. The proteomic approach yielded relatively low numbers of differentially produced proteins, while the DNA microarray analyses generated gene expression profiles for over 2000 of the annotated genes of *L. lactis* IL1403. Preliminary results indicate that, indeed, random mutagenesis via UV- or chemical treatment of cells introduces more differences than site-directed mutagenesis. Of the differentially expressed genes identified by the transcriptome analyses, only in a few cases were the corresponding proteins concordantly differentially produced. This reflects the complexity of comparing gene expression and protein production levels at a global scale and is a major challenge for further improvements in comparative global genomics analyses. The methodology adopted in the “Express Fingerprints” project is a first step in using modern genomics techniques in health risk assessment for legislation of improved bacterial strains [99] (see also the treatise of this subject by Pederson et al. [123] elsewhere in this issue).

5. Genomic diversity analysis using comparative genome hybridization

Comparison by bioinformatics methods of the genomic content of LAB has been described in a number of studies (for a recent overview, see [100]). The elucidation of the genome sequence of *L. lactis* ssp. *lactis* MG1363 has allowed comparing the genomes of the two subspecies of *L. lactis* (see above). A powerful method to quickly determine the genomic content of a bacterial strain of which the genome sequence is not known is comparative genome hybridization (CGH). This technique implies performing DNA–DNA hybridizations using equal amounts of DNA of unknown sequence and that of a close relative of which the genome sequence has been determined, on DNA microarrays of the latter. Fluorescently labeled DNA of each strain is prepared using (random-) primed PCR to create complementary DNA fragments carrying either Cy3-dye (green) or Cy5-dye (red). In the case that both strains carry the same gene a similar signal is expected from the labeled cDNA fragments of that gene, and thus a ratio of the unknown over the sequenced strain of 1. The

absence of genes in the non-sequenced strain will result in lower or no hybridization signals (ratio $\ll 1$).

CGH is commonly used to determine the genomic composition and genome plasticity of bacteria [101–103] or to perform strain typing of unknown bacterial strains [104,105]. So far, only two studies describe the use of CGH to determine the genomic composition in LAB. Siezen et al. [100] used CGH for partitioning of a number of *Lactobacillus plantarum* strains on the basis of their genomic content, a result that will be dealt with elsewhere in this issue. A more methodological approach used the genome information of *L. lactis* strains MG1363 and IL1403 to correlate the similarity level to DNA microarray hybridization efficiency (see below).

Perfectly conserved genes in two strains should ideally exhibit identical hybridization behavior in CGH. Generally, DNA microarrays are used to describe the genomic composition of an organism in a binary manner: genes are either present or absent. Two studies try to correlate hybridization efficiency to the degree of similarity between genes [102,106]. In both, however, difficulties are reported in determining these correlations for a large number of genes. Establishing a correlation between less divergent gene sequences and hybridization signals is a challenging task to be tackled next.

To examine whether differential hybridization for a given gene is due to differences in gene similarity levels, DNA microarrays of *L. lactis* IL1403 were hybridized with a mixture of randomly Cy3- or Cy5-labeled DNA fragments from the *L. lactis* IL1403 and *L. lactis* MG1363 (Baerends, R.J.S., van Hijum, S.A.F.T., Karsens, H.A., De la Nava, J.G., Trelles, O., Kok, J. and Kuipers, O.P., unpublished). As both genome sequences are known, the competitive hybridization behavior of all IL1403/MG1363 genes represented on the *L. lactis* IL1403 DNA microarray could be studied. A clear positive correlation exists between gene similarity and the obtained ratio for genes with a similarity of 75% and more (Fig. 5). In Fig. 5 a number of additional *L. lactis* IL1403 genes that are absent in *L. lactis* MG1363 are shown. As expected, ratios lower than 1 were obtained for all of these genes. Conversely, for *tra* 905, present 16 times in *L. lactis* MG1363 and only once in *L. lactis* IL1403, a ratio $\gg 1$ was obtained. The correlation between similarity level and hybridization efficiency, however, does not apply for all genes, which becomes clear from the relatively large scatter in ratios at similarity levels below 75% (Fig. 5). A possible reason for these differences is that dilute probes on the slide lead to more highly fluctuating signals than concentrated probes.

5.1. Comparing gene expression profiles across organisms

Comparison of global gene expressions of two isogenic strains by hybridization to a non-isogenic DNA

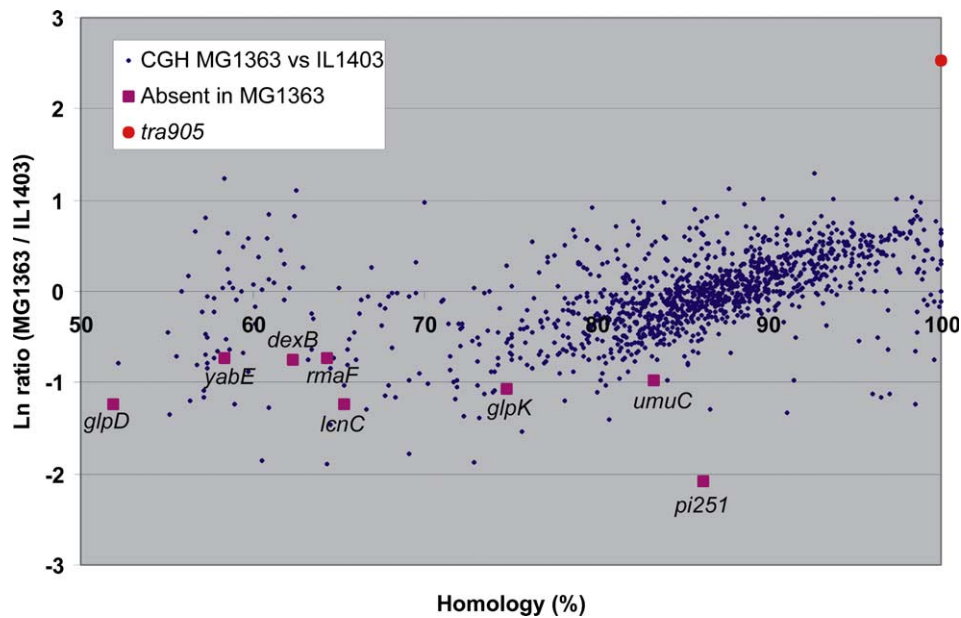


Fig. 5. Result of a CGH experiment of *L. lactis* MG1363 and *L. lactis* IL1403 DNAs hybridized on amplicon-based *L. lactis* IL1403 DNA microarrays [54,55]. These slides contain duplicate spots of amplicons corresponding to all IL1403 ORFs. The genomic DNA of both strains was labeled using the Invitrogen BioPrime DNA labeling system (Invitrogen, Carlsbad, CA). To exclude specific gene-with-dye interactions, dye-swaps were performed for each replicate. The background-subtracted data of the slides were not normalized. The average ratios (vertical axis: *L. lactis* MG1363 signals divided by the *L. lactis* IL1403 signals) were calculated for two replicates. Results are only shown for genes with reasonable coherent measurements (a coefficient of variance less than 50%). Gene similarity (horizontal axis: the percentage of conserved nucleotides) of *L. lactis* MG1363 ORFs to the sequences of *L. lactis* IL1403 amplicons on the DNA microarray was determined by FASTA (<http://www.ebi.ac.uk/fasta33>). Pink squares: genes that are absent in *L. lactis* MG1363, as determined by similarity searches both on the DNA (BLASTN) and protein (TBLASTX) level at a cut-off of 10^{-30} . Red closed circle: *tra905*, a gene which is present in 16 copies in *L. lactis* MG1363 and in 1 copy in *L. lactis* IL1403.

microarray is relatively straightforward as the differences in hybridization behavior of the two cDNAs on the non-isogenic probes on the slide are the same. Using the same methodology as in the “Express Fingerprints” project (see above), heterologous transcriptome comparisons, in which labeled cDNAs of two non-isogenic strains are mixed and compared on a DNA microarray of one of the two, could be performed to profile global gene expression in new strains [99]. Heterologous transcriptome comparisons open exciting possibilities of determining global expression patterns of organisms of which the genome sequences are not (entirely) known. The approach adds a next level of complexity to the analysis namely in the determination of the contribution of mRNA abundance rather than sequence divergence to the observed “differential expression” of genes. There are several other factors that add to the complexity such as mRNA-specific differences between two organisms in RNA isolation, labeling and hybridization. To more reliably determine the actual differential expression, CGH information can be used to correct for differences in gene similarity level [107]. However, the labeling methods for CGH and cDNA synthesis should be similar to rule out differences between CGH and transcriptome results due to labeling methods.

6. Global analyses of *L. lactis* – proteomics

The proteome, the total protein complement of a genome, is in most cases investigated by combining 2D gel electrophoresis and protein identification by peptide mass spectrometry and peptide sequencing using an MS–MS approach [108]. For proteome analysis the nucleotide sequence of a genome has to be known to be able to identify all proteins and products thereof, although current high-throughput proteome analyses using MS–MS identification platforms generate peptides masses and peptide sequences that can be directly compared to the ever-increasing protein databases. A full proteomics approach to understanding cell function would also assess the activities and modifications of all proteins as well as their cellular localizations, their interactions and their occurrence in (protein) complexes. The latter level of proteomics, called functional proteomics, has not been implemented yet in a high-throughput comprehensive way for *L. lactis* at the time of writing this review.

Like transcriptome analysis, proteomics studies can generate quantitative measures, in this case of protein levels. Combining transcriptomics and proteomics data one can determine whether regulation takes place at the transcriptional or the post-transcriptional level.

Several proteomics studies on different lactococcal strains have been performed and will be described below. A more comprehensive treatise of the subject, dealing *a.o.* with the purine stimulon and with proteins induced upon heat/cold shock or as a result of a shift to low pH can be found in two recent papers [Kilstrup, M., Nielsen, A.K., Pulka-Amin, M.D. and Rasmussen, M.D., unpublished; 109].

6.1. Generation of *L. lactis* proteome reference maps

A number of proteomics studies aimed at identifying all proteins, and in that way generating reference maps, of *L. lactis* strains grown under specific conditions. The first reference map was a reference 2D polyacrylamide gel (pH range 4–7) containing radioactively labeled proteins ($[^{35}\text{S}]$ -methionine) of *L. lactis* MG1363 growing exponentially in chemically defined glucose-SA medium [110]. Only a modest number of proteins, among which the heat-shock proteins GroEL and DnaK and a few glycolytic enzymes, were identified using Western blotting and N-terminal protein sequencing. The map was overlaid with a grid for easy mapping purposes. After growth of *L. lactis* NCDO763, a strain from the NCDO712 family described above and, thus, a relative of *L. lactis* MG1363, in chemically defined CDM medium, proteins were electrophoretically separated by size and *pI* (pH range of 4–7). Through silver staining 450 proteins could be detected, 15 of which were identified by peptide mass fingerprinting [111].

Several studies have made use of *L. lactis* IL1403. A prediction of the distribution of all proteins putatively expressed from the *L. lactis* IL1403 genome showed that 56% of the proteins make up the acidic sub-proteome (all proteins with a *pI* between 3.4 and 7) and 43% constitute the alkaline sub-proteome (*pI* > 7) [112]. Of the alkaline proteome less than 10% has a *pI* between 7 and 9 [109]. The cytosolic acidic proteome of *L. lactis* IL1403 has been examined after growth of the strain in glucose-M17 [112]. Cytosolic proteins were separated in the pH ranges 4–7 and 4.5–5.5 and, subsequently, 239 different proteins were identified from 265 proteins spots. This result represents approximately 25% of the predicted acid proteome and has been made publicly available in an *L. lactis* 2D database (http://genome.jouy.inra.fr/2dlactis/eng/index_eng.htm). All enzymes of the glycolytic pathway were among the highly expressed proteins.

Codons of abundant transfer RNAs are overrepresented in highly expressed genes [113]. A codon adaptation index (CAI; between 0 and 1), which takes into account the overall codon bias of the genome, is a good predictor for the likelihood of identification of a protein in 2D protein analyses: all predicted proteins in the acidic proteome of *L. lactis* IL1403 with a CAI value above 0.8 were identified while 75% and 54%, respec-

tively, of the proteins with CAI values of 0.7–0.8 and 0.6–0.7 were identified in the same study [112]. Of the 2049 ORFs of *L. lactis* IL1403 examined, the genes for the translation machinery and for glycolytic enzymes have the highest CAI values. The complete set of glycolytic proteins was found to account for more than 20% of the soluble acidic proteins, showing just how important glycolysis is to this fermentative bacterium.

The reference map of the acidic sub-proteome of *L. lactis* IL1403 was compared with that of *L. lactis* NCDO763 grown under similar conditions [112]. Out of 400 spots only 61 protein spots had a similar 2D-gel location for both strains. Sixty-three of 100 protein spots that did not have exact matches were shown to map at slightly different positions, most likely as a consequence of amino acid changes between the corresponding proteins in both subspecies. Twenty-six proteins of *L. lactis* NCDO763, including the enzymes for lactose assimilation, were not present in *L. lactis* IL1403, which agrees with the fact that *L. lactis* IL1403 is Lac^- due to the loss of the lactose plasmid. A comparison of the cytosolic proteomes (pH 4–7) of *L. lactis* NCDO763 grown on glucose or lactose showed an induction of the proteins encoded by the *lac* operon and the galactose operon on the latter sugar. A reduction of expression of the enzymes for the initial steps of pyrimidine synthesis was observed during growth in lactose, indicating that a link exists between the lactose pathway and pyrimidine nucleotides metabolism [112].

The introduction of a number of technical improvements has recently allowed analyzing the alkaline proteome of *L. lactis*: all proteins with an isoelectric point above 7 [114]. The alkaline proteome makes up almost 39% of all putatively expressed proteins of *L. lactis* IL1403. Of 200 visible protein spots, 153 could be identified as being products of 85 different genes: several proteins are present as (multiple) isoforms. These and other data have been stored in an online Proteome Database of *L. lactis* named DynaProt2D (<http://www.wzw.tum.de/proteomik/lactis/>) that has recently been published [115]. The database contains 2D reference maps of which the identified spots are clickable and linked to information such as molecular weight, isoelectric point, codon adaptation index, grand average of hydropathicity and cellular localization. The database allows online addition of gels and protein information.

6.2. Quantitative proteomics

Three studies employ quantitative proteomics: the determination of the fold-difference in expression of proteins in a cell as a consequence of certain perturbations. In fact, the study of Kilstrup et al. [110] described above is an example of quantitative proteomics. Frees et al. [116] followed protein expression with a $[^{35}\text{S}]$ -methio-

nine label after shifting of exponentially growing cells of *L. lactis* MG1363 to a low pH (pH 4.5 or 5.5). Over 20 proteins that were induced at pH 4.5 were identified. Among these were proteins such as alkyl hydroperoxide reductase and superoxide dismutase, which are involved in oxidative stress tolerance, and the heat shock proteins GrpE and ClpP. Using the same labeling method, Beyer et al. [117] compared protein patterns of cells of *L. lactis* MG1363 growing exponentially in chemically-defined glucose-SA medium with those during purine starvation (purine depletion of a *pur* mutant or decoynine addition). Three stimulons could be identified: (i) Psu (purine starvation up-regulated), containing the enzymes of the purine biosynthesis pathway and those of the C1 units required for purine biosynthesis. All genes were preceded by the so-called Pur-box, indicating that Psu-stimulon members are part of the PurR regulon; (ii) Psd (purine starvation down-regulated) encompassed proteins involved in energy metabolism, protein synthesis or GTPase function; (iii) Dcu (decoynine up-regulated) contained enzymes of the glycolytic pathway and factors involved in translation elongation.

Recently, Palmfeldt et al. [118] used quantification of spots after scanning of the 2D gel to determine the acidic proteomes of *L. lactis* cells growing anaerobically on glucose or maltose, and of resting cells metabolizing maltose. Under the latter condition the cells did not grow due to lack of an amino acid source. The results of the proteomic studies were coupled to the physiology of the cells. The maltose-metabolizing cells performed a mixed acid fermentation; under the other two conditions only lactate was produced. The cells growing on maltose expressed more proteins than those growing on glucose but only a limited number of these proteins had a direct relation to maltose fermentation. The maltose consumption rate in resting cells was approximately 3-fold lower than that of the cells actively growing on this sugar. As the enzyme levels under both circumstances were similar, this would indicate that regulation of maltose consumption rate and the product formation pattern take place at the level of enzyme activity rather than at the level of protein production. During amino acid starvation, in the presence of an energy and carbon source (maltose), only little protein degradation or synthesis takes place up to at least 15 h of starvation, while the cells stayed metabolically active [118].

The application to proteomics studies of *L. lactis* of new developments in the field, such as differential 2D fluorescence gel electrophoresis, in which proteins from different samples are labeled with different fluorescent dyes and quantitatively compared, unravelling of complex protein sample using multiplex MS–MS approaches in combination with high-capacity computing, as well as the developments in the protein-chip technology, will undoubtedly advance our understanding of this important bacterium.

7. Concluding remarks

The establishment of brute-force DNA cloning and sequencing techniques in the mid nineties of the last century has allowed the development of a novel field of genomics research. Concomitant advances in high-throughput robotic and computational techniques, and in bioinformatics research related to both the comprehensive analyses of large data sets and database management, now allow performing proteomics and transcriptomics research. The availability of the nucleotide sequence of *L. lactis* permits to apply these technologies to this economically important bacterium. A first brief comparison of the genome sequences of the two sequenced strains of *L. lactis*, IL1403 and MG1363, already revealed interesting differences between the two and a more thorough look at the nucleotide sequences will undoubtedly uncover more, especially with respect to their evolutionary relationship, the adaptation of the strains to a milk environment and the differences in their metabolic potential. The fact that we know or have a relatively good estimation of the functioning of only roughly 75% of the genes of either strain shows that a lot still needs to be learned. Among these genes with unknown function may be several that are of importance to, e.g., alternative metabolic routes or flavour development in products. The study of glucose metabolism described above (Pool, W.A., Neves, A.R., Kok, J., Santos, H. and Kuipers, O.P., unpublished) revealed a novel function of an annotated but poorly characterized PEP-PTS system (*ptcBAC*). The study on nisin resistance development has pointed to the involvement of a number of uncharacterized proteins in this trait in *L. lactis* and has extended this observation to similar proteins in a number of other gram-positive bacteria. One of the most important aspects of using transcriptomics analyses is the possibility it offers to look at entire gene regulatory networks that are operative in the organism under study. As a first step, the in silico analysis of the regulator content of *L. lactis* proved to be very informative [119]. Comparing the differences in the transcriptomes of regulator mutants with those of the isogenic parent strains has enabled to define the regulons of a number of *L. lactis* regulators. Some of these appeared to be very specific, controlling a relatively small number of genes. On the other hand, the pleiotropic regulator CodY of *L. lactis* regulates a wide range of genes and operons. By combining the DNA microarray results with bioinformatics techniques it was possible to identify the operator of the CodY regulator, which had been elusive until then. It will be extremely interesting to apply similar approaches to the study of the various responses of *L. lactis* to different stress conditions as, e.g., low pH, high lactate, – salt or temperature or, conversely, low temperature, different types of starvation, protein overproduction and – secretion, etc. As gene

regulation is a spatio-temporal process, especially the analysis of the response in time of the cells to changed internal (mutations) or external conditions will be highly clarifying. Clustering methods allow grouping of genes based on their temporal expression profiles and have already been used to analyze the mRNA pools of *L. lactis* cells in the different stages of growth in batch culture (Zomer, A.L., and Lulko, A.T., unpublished).

Other examples of the use of (oligonucleotide-based) DNA microarrays are in the field of strain typing and characterization or in the quick analysis through CGH of the genetic (metabolic) potential of new strains. Such techniques could be used to identify and isolate novel strains through non-invasive (non-recombinant) techniques: using the fundamental insights of the kind of genes that should be present for a certain phenotype, or the level of expression of such genes, industry could examine their strain collections for just those parameters. Ultimately, with the pace of high-throughput nucleotide sequencing still increasing rapidly, and prices going down concurrently, the possibility of “1000-dollar” genomes may become a reality soon. Again, this would give an enormous boost to both fundamental and applied research, as it would permit sequencing strains from the wide assortment of natural isolates, collected already or still present in the uncharacterized flora in local traditional fermentation products around the world. These strains represent a huge gene pool and potential for new, that is, not yet exploited fermentation capacity for feeding the world’s population with healthy products.

Another important application of DNA microarrays, in combination with proteomics studies, is their use for the assessment of “substantial equivalence” of strains. Using both technologies, the natural window of expression of all genes and the presence of all proteins of wild-type industrial strains could be determined. The same procedure could be used to define the proteome and transcriptome of genetically modified strains or novel “natural” strains that are to be used in (large-scale industrial) application but have not been used as such yet. In this way, a measure of acceptability for use (in foods) of the new strains could be derived, while genes/proteins of which the expression levels are outside the natural window could be examined with respect to function and possible consumer risk.

DNA microarrays could also be very helpful in the analysis of cell–cell interaction, being it the interaction between cells of the different bacterial strains or species in starter cultures, between starter strain and food spoilage or pathogenic bacteria, or even of the food bacteria with the cells of the consumer. Studies along these lines should uncover important aspects of microbial ecology, including bacteria–bacteriophage interactions, and could reveal entirely novel ways of fighting off unwanted bacteria in foods or elsewhere. The global study of the

communication between, especially, probiotic LAB strains and the (intestinal) cells of the consumer might reveal, e.g., important aspects of how these interactions could be enhanced to improve the probiotic nature of the bacteria.

Huge progress is also reported on the application of high-throughput protein analysis and identification technology. Above we have seen the examples of proteomic studies that aimed, first of all, to establish protein reference maps of *L. lactis* and detail the changes in protein profiles in response to changes in environmental conditions. The application of proteomics tools in the study of *L. lactis* will be of eminent importance for the understanding of regulation that takes place at the protein level (protein modification by degradation, phosphorylation, prenylation, etc.) and that will not be recorded using transcriptomics. This field of functional proteomics, which also includes the study of the cellular localization of protein (complexes) using, e.g., fusions of the proteins of interest with epitopes or variants of the Green Fluorescent Protein, and of protein–protein interactions (recorded using, e.g., various formats of yeast- or bacterial two- or three-hybrid systems) will undoubtedly take off soon. The studies presented above have been performed after growth of cells of a single species under laboratory conditions. Recently, Gagnaire et al. [120] produced a reference map of a mixed bacterial population (*Lb. helveticus*, *Lb. delbrueckii* ssp. *lactis* and *S. thermophilus*) and identified proteins released into Emmental cheese after lysis of the LAB. The analysis showed that some peptidases from *Lb. helveticus* and *S. thermophilus* are released into the cheese and, thus, likely contribute to peptide degradation. The release of bacterial enzymes in the curd and the subsequent documentation of casein degradation, both examined with proteomics techniques, will be the next way to go for the more applied (industrial) directions in product (quality) research.

L. lactis is a relatively simple bacterium fully geared towards rapidly degrading lactose via a homofermentative pathway. It has a small genome and relatively few paralogous genes. Tools to study it biochemically, physiologically, and genetically and to grow it in small or large quantities under very well-defined conditions in synthetic media are all in place. This makes the organism ideally suited as a model in a systems biology approach of understanding the intricacies of bacterial life. Systems biology aims at understanding the emergent properties of a biological system such as a bacterial cell, ideally as part of an entire (microbial) community, through examining the interactions of all genes, proteins, other macromolecules, metabolites, and the environment. Although a lot of hurdles still have to be taken, especially with respect to getting together a consortium diverse enough to include geneticists and microbiologists, mathematicians, biostatisticians, control

theorists, (bio)chemists, computer scientists and system engineers, it would be an exciting new way to go in research on *L. lactis*. We are confident that such a thorough analysis, involving transcriptomics, proteomics and metabolomics research, backed up by state-of-the-art bioinformatics and fundamental biological knowledge, would not only lead to many novel fundamental insights into *Lactococcus* growth and survival, but will also deliver tools and know-how for novel practical applications.

Acknowledgments

We are grateful to all our colleagues who provided material prior to publication. The authors wish to specifically acknowledge the following members of the MG1363 sequencing consortium (Udo Wegmann, Mary O'Connell-Motherway, Alexander Goesmann, Anne de Jong, Mike Gasson, Claire Shearman and Douwe van Sinderen) without whom the writing of this literature overview would not have been possible. We would like to thank Richard J.S. Baerends for critical reading of the manuscript. Part of the work described here was supported by the research programme "Quality of Life and Management of Living Resources" of the European Community, grants QLK1-CT-2000-01376 (Nutracells) and QLK3-2001-010473 (Express-Fingerprints).

References

- [1] Hols, P., Kleerebezem, M., Schanck, A.N., Ferain, T., Hugenholtz, J., Delcour, J. and de Vos, W.M. (1999) Conversion of *Lactococcus lactis* from homolactic to homoalanine fermentation through metabolic engineering. *Nat. Biotechnol.* 17, 588–592.
- [2] Sybesma, W., Starrenburg, M., Kleerebezem, M., Mierau, I., de Vos, W.M. and Hugenholtz, J. (2003) Increased production of folate by metabolic engineering of *Lactococcus lactis*. *Appl. Environ. Microbiol.* 69, 3069–3076.
- [3] Steidler, L., Hans, W., Schotte, L., Neirynck, S., Obermeier, F., Falk, W., Fiers, W. and Remaut, E. (2000) Treatment of murine colitis by *Lactococcus lactis* secreting interleukin-10. *Science* 289, 1352–1355.
- [4] Wells, J.M. and Annick Mercenier, A. (2003) Lactic acid bacteria as mucosal delivery vehicles. In: *Genetics of Lactic Acid Bacteria* (Wood, B. and Warner, P.J., Eds.), pp. 261–290. Kluwer Academic, New York.
- [5] Hanniffy, S., Wiedermann, U., Repa, A., Mercenier, A., Daniel, C., Fioramonti, J., Tlaskolova, H., Kozakova, H., Israelsen, H., Madsen, S., Vrang, A., Hols, P., Delcour, J., Bron, P., Kleerebezem, M. and Wells, J. (2004) Potential and opportunities for use of recombinant lactic acid bacteria in human health. *Adv. Appl. Microbiol.* 56, 1–64.
- [6] de Ruyter, P.G., Kuipers, O.P. and de Vos, W.M. (1996) Controlled gene expression systems for *Lactococcus lactis* with the food-grade inducer nisin. *Appl. Environ. Microbiol.* 62, 3662–3667.
- [7] Kunji, E.R.S., Slotboom, D.-J. and Poolman, B. (2003) *Lactococcus lactis* as host for overproduction of functional membrane proteins. *Biochim. Biophys. Acta* 1610, 97–108.
- [8] O'Brien, L., Kerrigan, S.W., Kaw, G., Hogan, M., Penades, J., Litt, D., Fitzgerald, D.J., Foster, T.J. and Cox, D. (2002) Multiple mechanisms for the activation of human platelet aggregation by *Staphylococcus aureus*: roles for the clumping factors ClfA and ClfB, the serine-aspartate repeat protein SdrE and protein A. *Mol. Microbiol.* 44, 1033–1044.
- [9] Campo, N., Dias, M.J., Daveran-Mingot, M.-L., Ritzenthaler, P. and Le Bourgeois, P. (2002) Genome plasticity in *Lactococcus lactis*. *Antonie van Leeuwenhoek* 82, 123–132.
- [10] Le Bourgeois, P., Lautier, M., van den Berghe, L., Gasson, M.J. and Ritzenthaler, P. (1995) Physical and genetic map of the *Lactococcus lactis* subsp. *cremoris* MG1363 chromosome: comparison with that of *Lactococcus lactis* subsp. *lactis* IL1403 reveals a large genome inversion. *J. Bacteriol.* 177, 2840–2850.
- [11] Davidson, B.E., Kordia, N., Baseggio, N., Lim, A., Dobos, M. and Hillier, A.J. (1995) Genomic organization of *Lactococci* Genetics of Streptococci, Enterococci, Enterococci and Lactococci. In: *Developmental Biology Standard* (Ferretti, J.J., Gilmore, M.S., Klaenhammer, T.R. and Brown, F., Eds.), Vol. 85, pp. 411–422. Karger, Basel.
- [12] Davidson, B.E., Kordia, N., Dobos, M. and Hillier, A.J. (1996) Genomic organization of lactic acid bacteria. In: *Lactic Acid Bacteria: Genetics, Metabolism and Applications* (Venema, G., Huis in't Veld, J.H.J. and Hugenholtz, J., Eds.), pp. 65–87. Kluwer Academic Publishers, Dordrecht.
- [13] McKay, L.L., Baldwin, K.A. and Efstathiou, J.D. (1976) Transductional evidence for plasmid linkage of lactose metabolism in *Streptococcus lactis* C2. *Appl. Environ. Microbiol.* 32, 45–52.
- [14] Gasson, M.J. and Davies, F.L. (1980) High frequency conjugation associated with *Streptococcus lactis* donor cell aggregation. *J. Bacteriol.* 133, 1260–1264.
- [15] Davies, F.L., Underwood, H.M. and Gasson, M.J. (1981) The value of plasmid profile for strain identification in lactic streptococci and the relationship between *Streptococcus lactis* 712, ML3 and C2. *J. Appl. Bacteriol.* 51, 325–337.
- [16] Walsh, P.M. and McKay, L.L. (1981) Recombinant plasmid associated with cell aggregation and high-frequency conjugation in *Streptococcus lactis* ML3. *J. Bacteriol.* 146, 937–944.
- [17] Tynkkynen, S., Buist, G., Kunji, E.R.S., Kok, J., Poolman, B., Venema, G. and Haandrikman, A.J. (1993) Genetic and biochemical characterization of the oligopeptide transport system of *Lactococcus lactis*. *J. Bacteriol.* 175, 7523–7532.
- [18] Yu, W., Gillies, K., Kondo, J.K., Broadbent, J.R. and McKay, L.L. (1996) Loss of plasmid-mediated oligopeptide transport system in lactococci: another reason for slow milk coagulation. *Plasmid* 35, 145–155.
- [19] Le Bourgeois, P., Daveran-Mingot, M.L. and Ritzenthaler, P. (2000) Genome plasticity among related *Lactococcus* strains: identification of genetic events associated with macrorestriction polymorphisms. *J. Bacteriol.* 182, 2481–2491.
- [20] Daveran-Mingot, M.-L., Campo, N., Ritzenthaler, P. and Le Bourgeois, P. (1998) A natural large chromosomal inversion in *Lactococcus lactis* is mediated by homologous recombination between two insertion sequences. *J. Bacteriol.* 180, 4834–4842.
- [21] Bolotin, A., Wincker, P., Mauger, S., Jaillon, O., Malarme, K., Weissenbach, J., Ehrlich, S.D. and Sorokin, A. (2001) The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* ssp. *lactis* IL1403. *Genome Res.* 11, 731–753.
- [22] Klaenhammer, T., Altermann, E., Arigoni, F., Bolotin, A., Breidt, F., Broadbent, J., Cano, R., Chaillou, S., Deutscher, J., Gasson, M., van de Guchte, M., Guzzo, J., Hartke, A., Hawkins, T., Hols, P., Hutkins, R., Kleerebezem, M., Kok, J., Kuipers, O., Lubbers, M., Maguin, E., McKay, L., Mills, D., Nauta, A., Overbeek, R., Pel, H., Pridmore, D., Saier, M., van Sinderen, D., Sorokin, A., Steele, J., O'Sullivan, D., de Vos, W., Weimer, B., Zagorec, M. and Siezen, R. (2002) Discovering

- lactic acid bacteria by genomics. *Antonie van Leeuwenhoek* 82, 29–58.
- [23] Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410.
- [24] Gasson, M.J., Swindell, S., Maeda, S. and Dodd, H.M. (1992) Molecular rearrangement of lactose plasmid DNA associated with high-frequency transfer and cell aggregation in *Lactococcus lactis* 712. *Mol. Microbiol.* 6, 3213–3223.
- [25] Gasson, M.J., Godon, J.-J., Pillidge, C.J., Eaton, T.J., Jury, K.L. and Shearman, C.A. (1995) Characterization and exploitation of conjugation in *Lactococcus lactis*. *Int. Dairy J.* 5, 757–762.
- [26] Mierau, I., Tan, P.S., Haandrikman, A.J., Mayo, B., Kok, J., Leenhouts, K.J., Konings, W.N. and Venema, G. (1993) Cloning and sequencing of the gene for a lactococcal endopeptidase, an enzyme with sequence similarity to mammalian enkephalinase. *J. Bacteriol.* 175, 2087–2096.
- [27] Rauch, P.J.G. and de Vos, W.M. (1992) Characterisation of the novel nisin-sucrose conjugative transposon Tn5276 and its insertion in *Lactococcus lactis*. *J. Bacteriol.* 174, 1280–1287.
- [28] Bolotin, A., Quinquis, B., Renault, P., Sorokin, A., Ehrlich, S.D., Kulakauskas, S., Lapidus, A., Goltsman, E., Mazur, M., Pusch, G.D., Fonstein, M., Overbeek, R., Kyprides, N., Purnelle, B., Prozzi, D., Ngui, K., Masuy, D., Hancy, F., Burteau, S., Boutry, M., Delcour, J., Goffeau, A. and Hols, P. (2004) Complete sequence and comparative genome analysis of the dairy bacterium *Streptococcus thermophilus*. *Nat. Biotechnol.* 22, 1554–1558.
- [29] Pavlovic, G., Burrus, V., Gintz, B., Decaris, B. and Guedon, G. (2004) Evolution of genomic islands by deletion and tandem accretion by site-specific recombination: ICES_{St}-related elements from *Streptococcus thermophilus*. *Microbiology* 150, 759–774.
- [30] Ochman, H., Elwyn, S. and Moran, N.A. (1999) Calibrating bacterial evolution. *Proc. Natl. Acad. Sci. USA* 96, 12638–12643.
- [31] Naumoff, D.G., Xu, Y., Glansdorff, N. and Labedan, B. (2004) Retrieving sequences of enzymes experimentally characterized but erroneously annotated: the case of the putrescine carbamoyltransferase. *BMC Genomics* 5, 52.
- [32] Garvey, P., Rince, A., Hill, C. and Fitzgerald, G.F. (1997) Identification of a *recA* homolog (*recALP*) on the conjugative lactococcal phage resistance plasmid pNP40: evidence of a role for chromosomally encoded *recAL* in abortive infection. *Appl. Environ. Microbiol.* 63, 1244–1251.
- [33] Romero, D.A. and Klaenhammer, T.R. (1993) Transposable elements in Lactococci: a review. *J. Dairy Sci.* 76, 1–19.
- [34] Schneider, D. and Lenski, R.E. (2004) Dynamics of insertion sequence elements during experimental evolution of bacteria. *Res. Microbiol.* 155, 319–327.
- [35] de Visser, J.A., Akkermans, A.D., Hoekstra, R.F. and de Vos, W.M. (2004) Insertion-sequence-mediated mutations isolated during adaptation to growth and starvation in *Lactococcus lactis*. *Genetics* 168, 1145–1157.
- [36] Bongers, R.S., Hoefnagel, M.H., Starrenburg, M.J., Siemerink, M.A., Arends, J.G., Hugenholtz, J. and Kleerebezem, M. (2003) IS981-mediated adaptive evolution recovers lactate production by *ldhB* transcription activation in a lactate dehydrogenase-deficient strain of *Lactococcus lactis*. *J. Bacteriol.* 185, 4499–4507.
- [37] Even, S., Lindley, N.D. and Coccain-Bousquet, M. (2001) Molecular physiology of sugar catabolism in *Lactococcus lactis* IL1403. *J. Bacteriol.* 183, 3817–3824.
- [38] Even, S., Lindley, N.D., Loubière, P. and Coccain-Bousquet, M. (2002) Dynamic response of catabolic pathways to autoacidification in *Lactococcus lactis*: transcript profiling and stability in relation to metabolic and energetic constraints. *Mol. Microbiol.* 45, 1143–1152.
- [39] Even, S., Lindley, N.D. and Coccain-Bousquet, M. (2003) Transcriptional, translational and metabolic regulation of glycolysis in *Lactococcus lactis* subsp. *cremoris* MG 1363 grown in continuous acidic cultures. *Microbiology* 149, 1935–1944.
- [40] Fontaine, L., Even, S., Soucaille, P., Lindley, N.D. and Coccain-Bousquet, M. (2001) Transcript quantification based on chemical labeling of RNA associated with fluorescent detection. *Anal. Biochem.* 298, 246–252.
- [41] Bremer, H. and Dennis, P.P. (1987) Modulation of chemical composition and other parameters of the cell by growth rate. In: *Escherichia Coli and Salmonella Typhimurium. Cellular and Molecular Biology* (Neidhardt, F.C., Ingraham, J.L., Low, K.B., Magasanik, B., Schaechter, M. and Umberger, H.E., Eds.), pp. 1527–1542. American Society for Microbiology, Washington, DC.
- [42] Hansen, M.C., Nielsen, A.K., Molin, S., Hammer, K. and Kistrup, M. (2001) Changes in rRNA levels during stress invalidates results from mRNA blotting: fluorescence in situ rRNA hybridization permits renormalization for estimation of cellular mRNA levels. *J. Bacteriol.* 183, 4747–4751.
- [43] Neidhardt, F.C. (2004) Chemical composition of *Escherichia coli*. In: *Escherichia Coli and Salmonella Typhimurium. Cellular and Molecular Biology* (Neidhardt, F.C., Ingraham, J.L., Low, K.B., Magasanik, B., Schaechter, M. and Umberger, H.E., Eds.), pp. 3–6. American Society for Microbiology, Washington, DC.
- [44] Xie, Y., Chou, L., Cutler, A. and Weimer, B. (2004) DNA macroarray profiling of *Lactococcus lactis* subsp. *lactis* IL1403 gene expression during environmental stresses. *Appl. Environ. Microbiol.* 70, 6738–6747.
- [45] Rallu, F., Gruss, A., Ehrlich, S.D. and Maguin, E. (2000) Acid- and multistress-resistant mutants of *Lactococcus lactis*: identification of intracellular stress signals. *Mol. Microbiol.* 35, 517–528.
- [46] Obis, D., Guillot, A., Gripon, J.C., Renault, P., Bolotin, A. and Mistou, M.Y. (1999) Genetic and biochemical characterization of a high-affinity betaine uptake system (BusA) in *Lactococcus lactis* reveals a new functional organization within bacterial ABC transporters. *J. Bacteriol.* 181, 6238–6246.
- [47] Kok, J. and Buist, G. (2003) Genetics of proteolysis in *Lactococcus lactis*. In: *Genetics of Lactic Acid Bacteria* (Wood, B. and Warner, P.J., Eds.), pp. 189–223. Kluwer Academic, New York.
- [48] Guédon, E., Renault, P., Ehrlich, S.D. and Delorme, C. (2001) Transcriptional pattern of genes coding for the proteolytic system of *Lactococcus lactis* and evidence for coordinated regulation of key enzymes by peptide supply. *J. Bacteriol.* 183, 3614–3622.
- [49] Guédon, E., Serron, P., Ehrlich, S.D., Renault, P. and Delorme, C. (2001) Pleiotropic transcriptional repressor CodY senses the intracellular pool of branched-chain amino acids in *Lactococcus lactis*. *Mol. Microbiol.* 40, 1227–1239.
- [50] Petranovic, D., Guedon, E., Sperandio, B., Delorme, C., Ehrlich, D. and Renault, P. (2004) Intracellular effectors regulating the activity of the *Lactococcus lactis* CodY pleiotropic transcription regulator. *Mol. Microbiol.* 53, 613–621.
- [51] den Hengst, C.D., Curley, P., Larsen, R., Buist, G., van Sinderen, D., Kuipers, O.P. and Kok, J. (2005) Probing the direct interactions between CodY and the oppD promoter of *Lactococcus lactis*. *J. Bacteriol.* 187, 512–521.
- [52] Gajic, O. (2003) Relationships between MDR Proteins, Bacteriocin Production and Roteolysis in *Lactococcus lactis*. Thesis University of Groningen, Groningen, The Netherlands.
- [53] Marugg, J.D., Meijer, W., van Kranenburg, R., Laverman, P., Bruinenberg, P.G. and de Vos, W.M. (1995) Medium-dependent regulation of proteinase gene expression in *Lactococcus lactis*:

- control of transcription initiation by specific dipeptides. *J. Bacteriol.* 177, 2982–2989.
- [54] Kuipers, O.P., de Jong, A., Baerends, R.J.S., van Hijum, S.A.F.T., Zomer, A.L., den Hengst, C.D., Kramer, N.E., Buist, G. and Kok, J. (2002) Transcriptome analysis and related databases of *Lactococcus lactis*. *Antonie van Leeuwenhoek* 82, 113–122.
- [55] van Hijum, S.A.F.T., de Jong, A., Baerends, R.J.S., Karsens, H.A., Kramer, N.E., Larsen, R., Den Hengst, C.D., Albers, C.J., Kok, J. and Kuipers, O.P. (2005) A novel scheme to assess factors involved in the reproducibility of DNA-microarray data. *BMC Genomics* 6, 77.
- [56] Sanz, Y., Lanfermeijer, F.C., Renault, P., Bolotin, A., Konings, W.N. and Poolman, B. (2001) Genetic and functional characterization of *dpp* genes encoding a dipeptide transport system in *Lactococcus lactis*. *Arch. Microbiol.* 175, 334–343.
- [57] Bergara, F., Ibarra, C., Iwamasa, J., Patarroyo, J.C., Aguilera, R. and Marquez-Magana, L.M. (2003) CodY is a nutritional repressor of flagellar gene expression in *Bacillus subtilis*. *J. Bacteriol.* 185 (10), 3118–3126.
- [58] Débarbouillé, M., Gardan, R., Arnaud, M. and Rapoport, G. (1999) Role of BkdR, a transcriptional activator of the sigL-dependent isoleucine and valine degradation pathway in *Bacillus subtilis*. *J. Bacteriol.* 181, 2059–2066.
- [59] Fisher, S.H., Rohrer, K. and Ferson, A.E. (1996) Role of CodY in regulation of the *Bacillus subtilis* hut operon. *J. Bacteriol.* 178, 3779–3784.
- [60] Fisher, S.H. (1999) Regulation of nitrogen metabolism in *Bacillus subtilis*: vive la difference!. *Mol. Microbiol.* 32, 223–232.
- [61] Inaoka, T., Takahashi, K., Ohnishi-Kameyama, M., Yoshida, M. and Ochi, K. (2003) Guanine nucleotides guanosine 5'-diphosphate 3'-diphosphate and GTP co-operatively regulate the production of an antibiotic bacilysin in *Bacillus subtilis*. *J. Biol. Chem.* 278, 2169–2176.
- [62] Kim, H.J., Kim, S.I., Ratnayake-Lecamwasam, M., Tachikawa, K., Sonenshein, A.L. and Strauch, M. (2003) Complex regulation of the *Bacillus subtilis* aconitase gene. *J. Bacteriol.* 185, 1672–1680.
- [63] Molle, V., Nakaura, Y., Shivers, R.P., Yamaguchi, H., Losick, R., Fujita, Y. and Sonenshein, A.L. (2003) Additional targets of the *Bacillus subtilis* global regulator CodY identified by chromatin immunoprecipitation and genome-wide transcript analysis. *J. Bacteriol.* 185, 1911–1922.
- [64] Serror, P. and Sonenshein, A.L. (1996) CodY is required for nutritional repression of *Bacillus subtilis* genetic competence. *J. Bacteriol.* 178, 5910–5915.
- [65] Serror, P. and Sonenshein, A.L. (1996) Interaction of CodY, a novel *Bacillus subtilis* DNA-binding protein, with the *dpp* promoter region. *Mol. Microbiol.* 20, 843–852.
- [66] Slack, F.J., Serror, P., Joyce, E. and Sonenshein, A.L. (1995) A gene required for nutritional repression of the *Bacillus subtilis* dipeptide permease operon. *Mol. Microbiol.* 15, 689–702.
- [67] Wray Jr., L.V., Ferson, A.E. and Fisher, S.H. (1997) Expression of the *Bacillus subtilis* *ureABC* operon is controlled by multiple regulatory factors including CodY, GlnR, TnrA, and Spo0H. *J. Bacteriol.* 179, 5494–5501.
- [68] Ratnayake-Lecamwasam, M., Serror, P., Wong, K.W. and Sonenshein, A.L. (2001) 12 *Bacillus subtilis* CodY represses early-stationary-phase genes by sensing GTP levels. *Genes Dev.* 15, 1093–1103.
- [69] Shivers, R.P. and Sonenshein, A.L. (2004) Activation of the *Bacillus subtilis* global regulator CodY by direct interaction with branched-chain amino acids. *Mol. Microbiol.* 53, 599–611.
- [70] Burkovski, A. (2003) I do it my way: regulation of ammonium uptake and ammonium assimilation in *Corynebacterium glutamicum*. *Arch. Microbiol.* 179, 83–88.
- [71] Jakoby, M., Kramer, R. and Burkovski, A. (1999) Nitrogen regulation in *Corynebacterium glutamicum*: isolation of genes involved and biochemical characterization of corresponding proteins. *FEMS Microbiol. Letts.* 173, 303–310.
- [72] Jakoby, M., Nolden, L., Meier-Wagner, J., Kramer, R. and Burkovski, A. (2000) AmtR, a global repressor in the nitrogen regulation system of *Corynebacterium glutamicum*. *Mol. Microbiol.* 37, 964–977.
- [73] Magasanik, B. (1996) Regulation of nitrogen utilization. In: *Escherichia coli and Salmonella: Cellular and Molecular Biology* (Neidhardt, F.C., Curtiss III, R., Ingraham, J.L., Lin, B., Low, B., Magasanik, B., Reznikoff, W.S., Riley, M., Schaechter, M. and Umberger, H.E., Eds.), pp. 1344–1356. American Society for Microbiology, Washington, DC.
- [74] Reitzer, L. (2003) Nitrogen assimilation and global regulation in *Escherichia coli*. *Annu. Rev. Microbiol.* 57, 155–176.
- [75] Larsen, R. (2005) Transcriptional Control of Central Amino Acid Metabolism in *Lactococcus lactis*. Thesis University of Groningen, Groningen, The Netherlands.
- [76] Casiano-Colon, A. and Marquis, R.E. (1988) Role of the arginine deiminase system in protecting oral bacteria and an enzymatic basis for acid tolerance. *Appl. Environ. Microbiol.* 54, 1318–1324.
- [77] Crow, V.L. and Thomas, T.D. (1982) Arginine metabolism in lactic streptococci. *J. Bacteriol.* 150, 1024–1032.
- [78] Montel, M.C. and Champomier, M.C. (1987) Arginine catabolism in *Lactobacillus sake* isolated from meat. *Appl. Environ. Microbiol.* 53, 2683–2685.
- [79] Rallu, F., Gruss, A. and Maguin, E. (1996) *Lactococcus lactis* and stress. *Antonie van Leeuwenhoek* 70, 243–251.
- [80] Takahashi, N. and Yamada, T. (1999) Acid-induced acid tolerance and acidogenicity of non-mutans streptococci. *Oral. Microbiol. Immunol.* 14, 43–48.
- [81] Larsen, R., Buist, G., Kuipers, O.P. and Kok, J. (2004) ArgR and AhrC are both required for regulation of arginine metabolism in *Lactococcus lactis*. *J. Bacteriol.* 186, 1147–1157.
- [82] Song, H., Wang, H., Gigot, D., Dimova, D., Sakanyan, V., Glansdorff, N. and Charlier, D. (2002) Transcription regulation in thermophilic bacteria: high resolution contact probing of *Bacillus stearothermophilus* and *Thermotoga neapolitana* arginine repressor-operator interactions. *J. Mol. Biol.* 315, 255–274.
- [83] Maas, W.K. (1994) The arginine repressor of *Escherichia coli*. *Microbiol. Rev.* 58, 631–640.
- [84] Czaplewski, L.G., North, A.K., Smith, M.C., Baumberg, S. and Stockley, P.G. (1992) Purification and initial characterization of AhrC: the regulator of arginine metabolism genes in *Bacillus subtilis*. *Mol. Microbiol.* 6, 267–275.
- [85] Cunin, R., Eckhardt, T., Piette, J., Boyen, A., Pierard, A. and Glansdorff, N. (1983) Molecular basis for modulated regulation of gene expression in the arginine regulon of *Escherichia coli* K-12. *Nucl. Acids Res.* 11, 5007–5019.
- [86] Thijs, G., Marchal, K., Lescot, M., Rombauts, S., de Moor, B., Rouze, P. and Moreau, Y. (2002) A Gibbs sampling method to detect overrepresented motifs in the upstream regions of coexpressed genes. *J. Comput. Biol.* 9, 447–464.
- [87] Martinussen, J., Schallert, J., Andersen, B. and Hammer, K. (2001) The pyrimidine operon *pyrRBP-carA* from *Lactococcus lactis*. *J. Bacteriol.* 183, 2785–2794.
- [88] Larsen, R., Kok, J. and Kuipers, O.P. (2005) Interaction between ArgR and AhrC controls regulation of arginine metabolism in *Lactococcus lactis*. *J. Biol. Chem.* 280, 19319–19330.
- [89] Stambolian, D. (1988) Galactose and cataract. *Surv. Ophthalmol.* 32, 333–349.
- [90] Thompson, J., Chassy, B.M. and Egan, W. (1985) Lactose metabolism in *Streptococcus lactis*: studies with a mutant lacking glucokinase and mannose-phosphotransferase activities. *J. Bacteriol.* 162, 217–223.

- [91] Maeda, S. and Gasson, M.J. (1986) Cloning, expression and location of the *Streptococcus lactis* gene for phospho-beta-D-galactosidase. *J. Gen. Microbiol.* 132, 331–340.
- [92] Loo, C.Y., Mittrakul, K., Voss, I.B., Hughes, C.V. and Ganeshkumar, N. (2003) Involvement of an inducible fructose phosphotransferase operon in *Streptococcus gordonii* biofilm formation. *J. Bacteriol.* 185, 6241–6254.
- [93] Gaurivaud, P., Laigret, F., Garnier, M. and Bove, J.M. (2001) Characterization of FruR as a putative activator of the fructose operon of *Spiroplasma citri*. *FEMS Microbiol. Letts.* 198, 73–78.
- [94] Barrière, C., Veiga-da-Cunha, M., Pons, N., Guédon, E., van Hijum, S., Kok, J., Kuipers, O.P., Ehrlich, S.D. and Renault, P. (in press) Fructose utilization in Gram-positive bacteria: its regulator, signal and DNA-binding site. *J. Bacteriol.* 187, 3752–3761.
- [95] Kramer, N.E. (2005) Nisin Resistance in Gram-positive Bacteria. Thesis University of Groningen, Groningen, The Netherlands.
- [96] Breukink, E., Wiedemann, I., van Kraaij, C., Kuipers, O.P., Sahl, H.-G. and de Kruijff, B. (1999) Use of the cell wall precursor lipid II by a pore-forming peptide antibiotic. *Science* 286, 2361–2364.
- [97] Kramer, N.E., Smid, E.J., Kok, J., de Kruijff, B., Kuipers, O.P. and Breukink, E. (2004) Sensitivity or resistance of Gram-positive bacteria to nisin is not determined by the amount of the receptor lipid II. *FEMS Microbiol. Lett.* 239, 157–161.
- [98] Gravesen, A., Sorensen, K., Aarestrup, F.M. and Knochel, S. (2001) Spontaneous nisin-resistant *Listeria monocytogenes* mutants with increased expression of a putative penicillin-binding protein and their sensitivity to various antibiotics. *Microbial Drug Resist.* 7, 127–135.
- [99] Renault, P. (2002) Genetically modified lactic acid bacteria: applications to food or health and risk assessment. *Biochimie* 84, 1073–1087.
- [100] Siezen, R.J., van Enckevort, F.H., Kleerebezem, M. and Teusink, B. (2004) Genome data mining of lactic acid bacteria: the impact of bioinformatics. *Curr. Opin. Biotechnol.* 15, 105–115.
- [101] Joyce, E.A., Chan, K., Salama, N.R. and Falkow, S. (2002) Redefining bacterial populations: a post-genomic reformation. *Nat. Rev. Genet.* 3, 462–473.
- [102] Dobrindt, U., Agerer, F., Michaelis, K., Janka, A., Buchrieser, C., Samuelson, M., Svanborg, C., Gottschalk, G., Karch, H. and Hacker, J. (2003) Analysis of genome plasticity in pathogenic and commensal *Escherichia coli* isolates by use of DNA arrays. *J. Bacteriol.* 185, 1831–1840.
- [103] Dong, Y., Glasner, J.D., Blattner, F.R. and Triplett, E.W. (2001) Genomic interspecies microarray hybridization: rapid discovery of three thousand genes in the maize endophyte, *Klebsiella pneumoniae* 342, by microarray hybridization with *Escherichia coli* K-12 open reading frames. *Appl. Environ. Microbiol.* 67, 1911–1921.
- [104] Li, J., Helmerhorst, E.J., Leone, C.W., Troxler, R.F., Yaskell, T., Haffajee, A.D., Socransky, S.S. and Oppenheim, F.G. (2004) Identification of early microbial colonizers in human dental biofilm. *J. Appl. Microbiol.* 97, 1311–1318.
- [105] Stevens, K.A. and Jaykus, L.A. (2004) Direct detection of bacterial pathogens in representative dairy products using a combined bacterial concentration-PCR approach. *J. Appl. Microbiol.* 97, 1115–1122.
- [106] Costa, D.O., Yanai, G.M., Muto, N.H., Leite, D.B., de Souza, A.A., Coletta Filho, H.D., Machado, M.A. and Nunes, L.R. (2002) Competitive hybridization on spotted microarrays as a tool to conduct comparative genomic analyses of *Xylella fastidiosa* strains. *FEMS Microbiol. Letts.* 216, 15–21.
- [107] Myers, C.L., Dunham, M.J., Kung, S.Y. and Troyanskaya, O.G. (2004) Accurate detection of aneuploidies in array CGH and gene expression microarray data. *Bioinformatics* 20, 3533–3543.
- [108] Agaton, C., Uhlen, M. and Hober, S. (2004) Genome-based proteomics. *Electrophoresis* 25, 1280–1288.
- [109] Champomier-Verges, M.C., Maguin, E., Mistou, M.Y., Anglade, P. and Chich, J.F. (2002) Lactic acid bacteria and proteomics: current knowledge and perspectives. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* 771, 329–342.
- [110] Kistrup, M., Jacobsen, S., Hammer, K. and Vogensen, F.K. (1997) Induction of heat shock proteins DnaK, GroEL, and GroES by salt stress in *Lactococcus lactis*. *Appl. Environ. Microbiol.* 63, 1826–1837.
- [111] Anglade, P., Demey, E., Labas, V., Le Caer, J.P. and Chich, J.F. (2000) Towards a proteomic map of *Lactococcus lactis* NCDO 763. *Electrophoresis* 21, 2546–2549.
- [112] Guillot, A., Gitton, C., Anglade, P. and Mistou, M.Y. (2003) Proteomic analysis of *Lactococcus lactis*, a lactic acid bacterium. *Proteomics* 3, 337–354.
- [113] Sharp, P.M. and Li, W.H. (1987) The codon Adaptation Index – a measure of directional synonymous codon usage bias, and its potential applications. *Nucl. Acids Res.* 15, 1281–1295.
- [114] Drews, O., Reil, G., Parlar, H. and Gorg, A. (2004) Setting up standards and a reference map for the alkaline proteome of the Gram-positive bacterium *Lactococcus lactis*. *Proteomics* 4, 1293–1304.
- [115] Drews, O. and Gorg, A. (2005) DynaProt 2D: an advanced proteomic database for dynamic online access to proteomes and two-dimensional electrophoresis gels. *Nucl. Acids Res.* 33 (Database Issue), D583–D587.
- [116] Frees, D., Vogensen, F.K. and Ingmer, H. (2003) Identification of proteins induced at low pH in *Lactococcus lactis*. *Int. J. Food Microbiol.* 87, 293–300.
- [117] Beyer, N.H., Roepstorff, P., Hammer, K. and Kistrup, M. (2003) Proteome analysis of the purine stimulon from *Lactococcus lactis*. *Proteomics* 3, 786–797.
- [118] Palmfeldt, J., Levander, F., Hahn-Hagerdal, B. and James, P. (2004) Acidic proteome of growing and resting *Lactococcus lactis* metabolizing maltose. *Proteomics* 4, 3881–3898.
- [119] Guédon, E., Jamet, E. and Renault, P. (2002) Gene regulation in *Lactococcus lactis*: the gap between predicted and characterized regulators. *Antonie van Leeuwenhoek* 82, 93–112.
- [120] Gagnaire, V., Piot, M., Camier, B., Vissers, J.P., Jan, G. and Leonil, J. (2004) Survey of bacterial proteins released in cheese: a proteomic approach. *Int. J. Food Microbiol.* 94, 185–201.
- [121] van Hijum, S.A.F.T., Zomer, A.L., Kuipers, O.P. and Kok, J. (in press) Projector 2: contig mapping for efficient gap-closure of prokaryotic genome sequence assemblies. *Nucl. Acids Res.*
- [122] Kistrup, M., Hammer, K., Ruhdal-Jensen, P. and Martinussen, J. (2005) Nucleotide metabolism and its control in lactic acid bacteria. *FEMS Microbiol. Rev.*, doi:10.1016/j.femsre.2005.04.006, this issue.
- [123] Pederson, M.B., Iversen, S.L., Sørensen, K.I. and Johansen, E. (2005) The long and winding road from the research laboratory to industrial applications of lactic acid bacteria. *FEMS Microbiol. Rev.*, doi:10.1016/j.femsre.2005.04.001, this issue.